[Parties and Counsel Listed on Signature Pages]

UNITED STATES DISTRICT COURT

FOR THE NORTHERN DISTRICT OF CALIFORNIA

IN RE: SOCIAL MEDIA ADOLESCENT ADDICTION/PERSONAL INJURY PRODUCTS LIABILITY LITIGATION

This Document Relates To:

Breathitt County Board of Education v. Meta Platforms, Inc., et al.

Tucson Unified School District v. Meta Platforms, Inc., et al.

Charleston County School District v. Meta Platforms, Inc., et al.

Irvington Public Schools v. Meta Platforms, Inc., et al.

Dekalb County School District v. Meta Platforms, Inc., et al.

Board of Education of Harford County v. Meta Platforms, Inc., et al.

MDL No. 3047

Case No. 4:22-md-03047-YGR (PHK)

PLAINTIFFS' CORRECTED OMNIBUS OPPOSITION TO DEFENDANTS' MOTIONS FOR SUMMARY JUDGMENT

Judge: Hon. Yvonne Gonzalez Rogers
Magistrate Judge: Hon. Peter H. Kang

Date: January 26, 2026
Time: 8:00 AM
Place: Courtroom 1, 4th Floor

TABLE OF CONTENTS

TABLE OF AUTHORITIES

Cases

Statutes & Regulations

Other Authorities

I. INTRODUCTION AND "ROAD MAP"

American teens are facing an unprecedented mental health crisis, and Defendants' social media platforms are a substantial contributing factor. As U.S. Surgeon General Vivek H. Murthy stated, "it is time to require a surgeon general's warning label on social media platforms, stating that social media is associated with significant mental health harms for adolescents." Vivek H. Murthy, *It's Time to Require a Surgeon General's Warning on Social Media Platforms, N.Y. Times* (June 17, 2024). As the harmful effects of Defendants' platforms have mounted, the Plaintiff school districts ("Districts") have been forced to expand counseling, intervention, and mental health services, diverting limited resources from their core mission of education. Defendants attempt to minimize this reality, placing the word *"crisis"* in quotation marks, but the impact on the Districts is undeniable: they have been compelled to address student well-being, at the expense of their core mission to educate. These costs are not static; recent survey data and expert analyses demonstrate that the harms to student mental health—and the resulting burdens on school systems—are ongoing and will continue into the future.

Each Defendant claims it owed no duty to the Districts, that the harms were unforeseeable, or that federal immunity doctrines shield its inaction. The record proves otherwise. Defendants knew their platforms' deliberate exploitation of children's attention for profit was fueling a youth mental health crisis and destabilizing schools. Yet they concealed those facts and failed to warn either the Districts or students and their parents. Defendants' internal research and independent studies leave no doubt: the harms to youth mental health and school systems are measurable, persistent, and worsening—proof that these outcomes were not only foreseeable but inevitable under Defendants' business model.

Defendants' conduct has substantially and foreseeably interfered with the right of children to attend school in an environment conducive to learning and free from preventable psychological harm. The resulting harms from this public nuisance—a widespread mental-health crisis, disrupted classrooms, and the redirection of educational resources—are not isolated or accidental; they are the predictable outcome of corporate choices to maximize youth engagement and advertising revenue at the expense of the Districts and student well-being.

Defendants begin each of their motions by emphasizing the preexisting challenges school Districts already face—poverty, racism, community violence, or natural disasters—as though those realities

absolve them of responsibility. But the law rejects that defense, recognizing that vulnerability does not excuse misconduct; it magnifies its consequences. Every district endured COVID-19 and its disruptions, but that shared challenge does not diminish Defendants' accountability for the distinct harms their platforms imposed on already fragile educational systems.

This Omnibus Response provides the Court with a consolidated explanation of why Defendants' overlapping motions fail, highlighting the common factual and legal disputes that must be resolved by a finder of fact—disputes over duty, breach, causation, and damages that are the same for every District and cannot be resolved on summary judgment. In addition to this filing, each of the six bellwether Districts has submitted its own individual opposition applying their governing state law to their particular facts.

This filing is organized as follows:

Section II – Legal Standard: Establishes that under Rule 56, summary judgment is improper where, as here, material facts are in dispute or credibility determinations are required.

Section III(A) – Liability: Makes clear that Defendants owed a duty of care to act reasonably to prevent foreseeable harm to students and schools and breached that duty by 1) knowingly and substantially causing and exacerbating the youth mental-health crisis, which the Districts are forced to address, and 2) disrupting District operations and negatively impacting the school environment. The record also establishes that Defendants' conduct created and maintained a public nuisance by unreasonably interfering with the public rights to health, safety, and education.

Section III(B) – Causation: Explains that Defendants' "entanglement" arguments misconstrue both law and fact. Plaintiffs are not required to isolate a single cause of harm; concurrent causes are for the jury to resolve. And the record shows a direct nexus between Defendants' creation of a nuisance and their negligent practices, omissions, and failures to warn, and the harms suffered by the Districts.

Section III(C) – Past Damages: Details the substantial evidence showing the Districts have incurred compensable losses, including increased expenditures and diversion of school resources. Expert analyses from Dr. Ward and Mr. Klein confirm that the value of teacher and staff time diverted to address disruptions caused by Defendants' platforms represents a recognized and quantifiable injury under law.

Section III(D) – Future Damages: Establishes that the Districts' evidence of ongoing harm is concrete and non-speculative. Future damages and abatement remedies are recognized and recoverable

under the laws of each jurisdiction.

Collectively, the six District filings and this Omnibus Response show that Defendants' platforms have imposed real and ongoing harms on the Districts. The record—both shared and District-specific—contains ample evidence of the creation of a public nuisance, and of duty, breach, causation, and damages. Those factual disputes must be resolved by a jury, not on summary judgment.

II. LEGAL STANDARD

Summary judgment is appropriate only where there is no genuine dispute of material fact and the movant is entitled to judgment as a matter of law. *See* Fed. R. Civ. P. 56(a); *Celotex Corp. v. Catrett,* 477 U.S. 317, 322 (1986). A dispute is "genuine" if a reasonable jury could return a verdict for the nonmoving party, and "material" if it could affect the outcome of the case. *Anderson v. Liberty Lobby, Inc.*, 477 U.S. 242, 248 (1986). When evaluating a motion for summary judgment, the Court must construe the evidence and draw all reasonable inferences in favor of the non-movant. *Matsushita Elec. Indus. Co. v. Zenith Radio Corp.*, 475 U.S. 574 (1986); *Brookside Assocs. v. Rifkin,* 49 F.3d 490 (9th Cir. 1995); *Ogden v. Bumble Bee Foods, LLC*, 2014 WL 27527, at 6 (N.D. Cal. Jan. 2, 2014). The Court may not weigh evidence or make credibility determinations. *Masson v. New Yorker Mag., Inc.,* 501 U.S. 496, 520 (1991) (citing *Anderson*, 477 U.S. at 255); *Atari Interactive, Inc. v. Redbubble, Inc.*, 515 F. Supp. 3d 1089, 1099 (N.D. Cal. 2021). Because the Districts have presented ample evidence creating genuine disputes of material fact, summary judgment must be denied.

III. ARGUMENT

A. Genuine material disputes of fact concerning Defendants' liability preclude summary judgment.

1. Defendants are liable to the Districts for failing to warn about the risks of their platforms.

Defendants have fueled a youth mental health crisis by deliberately targeting school-aged children with addictive social media platforms designed to maximize engagement and Defendants' profit. They have also targeted schools, actively promoting student use of their platforms during the school day, while failing to warn the Districts, their students, and parents about the serious mental health risks posed by their platforms. The Districts have been left to bear the costs of the ensuing harms, including a compromised educational environment and expenditure and diversion of resources to address student distraction and

mental health problems.

Defendants now seek to avoid responsibility for these harms, contending they owed no duty to the Districts because the Districts were not direct users of their platforms. But the Court has already held that Defendants *do* owe a duty to the Districts, and Defendants offer no basis to revisit that legal determination. Defendants' attempt to relitigate this issue is as substantively baseless as it is procedurally improper. Defendants owe a duty to warn the Districts of the harms caused by their deliberately addictive platforms, as well as a duty to warn students and parents of the serious mental health risks their platforms pose. The Districts have presented substantial evidence demonstrating they were harmed as a result of Defendants' breach of those duties. Defendants cannot escape liability by disregarding the predictable impact of their conduct on the very schools left to confront and manage the crisis they created.

a) The Court's prior rulings establish Defendants owed a duty to the Districts.

This Court held that Defendants owe the Districts a duty of care grounded in foreseeability, the relationship between Defendants' conduct and the Districts' harms, and public policy. *In re Soc. Media Adolescent Addiction/Pers. Inj. Prods. Liab. Litig.*, 754 F. Supp. 3d 946, 972–78 (N.D. Cal. 2024) ("MTD Order"). In denying Defendants' motion to dismiss the Districts' negligence claims, the Court concluded that the Districts plausibly alleged a legal duty of care and expressly allowed their failure-to-warn theories to proceed. *Id.* The Court explained that "three fundamental considerations" guide the duty analysis across jurisdictions:[1] "(1) the relationship between the parties, in particular, the relationship between the defendant's conduct and the plaintiff's injury; (2) the foreseeability of the plaintiff's injury; and (3) public policy concerns." *Id.* at 973. The Court found that all three factors supported a duty of care. *Id.* at 973–78.

Defendants now ignore these settled determinations. Instead, Defendants mischaracterize the Court's ruling, incorrectly claiming it recognized merely a "general" duty, which excludes any duty to warn of "financial harms" from student platform use. *See, e.g.*, Defs.' Breathitt Mot. at 35. But the Court expressly rejected Defendants' effort to dismiss failure-to-warn theories concerning "known risks of

[1] Under Arizona law, foreseeability is not a factor in the duty analysis but the relationship between the parties and public policy concerns are factors. *Gipson v. Kasey*, 150 P.3d 228, 231–34 (2007)

addiction attendant to any platform features or as to platform construction in general," and held that Defendants owed the Districts a duty of care arising from their "intentional design choices to foster compulsive use in their minor users and a failure to warn thereof." MTD Order, 754 F. Supp. 3d at 963–64, 978. The Court further found that, "as [Defendants] became sophisticated in their targeting and capture of minor users as to instill compulsive use of the platforms . . . it was objectively reasonable to expect the alleged corollary injuries." *Id.* at 975.

This makes sense. The direct impact on the Districts includes financial harms, such as the expenditure and diversion of resources to combat student use of Defendants' platforms during the school day and to address the negative impacts of Defendants' platforms on students. These were not obscure, unknowable consequences of Defendants' actions, but rather a natural result of Defendants' choice to target schools and school-aged children and to design platforms promoting compulsive use, including during school hours. *See infra* at § III.A.2. The Districts should have been warned about these harms. Defendants' argument also ignores that if Defendants had warned the public of the potential negative impacts on students—even without warning the Districts—the Districts could have taken steps to prevent and mitigate the resulting financial harms they experienced.

Defendants cannot relitigate the Court's prior legal determinations at summary judgment. "[W]hen a court decides upon a rule of law, that decision should continue to govern the same issues in subsequent stages in the same case." *Doe 1 v. Mayorkas*, 530 F. Supp. 3d 893, 909 n.8 (N.D. Cal. 2021) (quoting *Musacchio v. United States*, 577 U.S. 237, 244–45 (2016)). The Court's findings remain controlling—and the record only strengthens them through evidence of Defendants' knowledge of youth engagement, school-day disruption, and the foreseeable burdens their conduct imposed on schools.

b) Defendants owe a duty to warn the Districts directly.

Defendants owe a duty to warn the Districts as parties directly and foreseeably harmed by their platforms' design and operation. Under well-established tort principles, a duty to warn arises when a defendant's conduct creates a foreseeable risk of harm to others. MTD Order, 754 F. Supp. 3d at 973 (duty arises where there is "sufficient juxtaposition of the parties in time and space to place the plaintiff in danger from the defendant's acts") (quoting *Quisenberry v. Huntington Ingalls Inc.*, 818 S.E.2d 805, 811 (Va. 2018)). This duty extends to all those the defendant should reasonably expect to be harmed by their

conduct. *See State of Maryland v. Exxon Mobil Corp.*, 406 F. Supp. 3d 420, 463 (D. Md. 2019) ("duty to warn extends . . . to third persons whom the supplier should expect to be endangered") (cleaned up).

Defendants' own internal research, marketing strategies, and school outreach confirm they knew their platform designs were fostering compulsive use among students and producing widespread consequences for schools. *See infra* §§ III.A.2.a.2-3, b.2-3, c.2-3, d.2-3 (discussing Defendants' targeting of young people and schools); III.A.2.a.4, b.4, c.4, d.4 (discussing Defendants' knowledge of harms). Internally, Meta researchers minced no words: "IG is a drug… We're basically pushers." Ex. 74 at 2152.[2] "Teens are hooked despite how it makes them feel. Instagram is addictive, and time-spend on platform is having a negative impact on mental health" Ex. 75 at 1428. Meanwhile, TikTok researched what "fuels usage and adoption in High Schools." Ex. 423 at 7326. Snap acknowledged that "the ones that have the Snapchat addiction have no room for anything else. Snap dominates their life." Ex. 922 at 4154. And YouTube—despite knowing it "is the most widely used internet Platform by US teens"—recognized "growing concerns that excessive screen time is stunting physical, emotional, and social development of teens." Ex. 772 at 7914.

By deliberately targeting school-aged children and embedding features that promote compulsive engagement, Defendants created and amplified foreseeable harms the Districts now must address every day. Having created and profited from the very risks that injured the Districts, Defendants cannot now escape liability by artificially narrowing the scope of their duty to warn by framing the Districts' claims as a "novel" attempt to impose a duty on "indirectly injured" non-users of their platforms. *See, e.g.,* Defs.' Breathitt Mot. at 34-35. This Court has already rejected that idea. It is not novel to hold Defendants liable for failing to warn those directly and foreseeably harmed by their conduct. *See* MTD Order, 754 F. Supp. 3d at 974 (holding "that 'the school Districts' injuries are sufficiently 'direct'" to impose a duty of care). The Court's prior holding is squarely aligned with a long line of cases recognizing that those who create foreseeable dangers must take reasonable steps to warn all who are likely to be affected—not merely end-users. *See, e.g.*, *Dozier Crane & Mach., Inc. v. Gibson*, 644 S.E.2d 333, 336 (Ga. App. 2007) ("The common-law duty imposed upon suppliers of chattels includes the duty to warn of foreseeable dangers

[2] All exhibits referenced in this brief are attached to the accompanying Declaration of Previn Warren.

arising from the reasonable use for which the product is intended and requires the exercise of reasonable care to inform third persons of the dangerous condition or of the facts which make the product likely to become dangerous."); *State of Maryland v. Exxon Mobil Corp.*, 406 F. Supp. 3d 420, 499 (D. Md. 2019) ("duty to warn extends not only to those for whose use the chattel is supplied but also to third persons whom the supplier should expect to be endangered by its use") (cleaned up); *New Jersey Dep't of Env't Prot. v. E.I. du Pont de Nemours & Co.*, 2021 WL 6144081, at *7 (D.N.J. Dec. 30, 2021) ("New Jersey courts have not limited the duty to warn to those with whom a defendant has a direct relationship"); *Williams v. Schneider Elec. USA, Inc.*, 2023 WL 4374514, at *4 (Ky. Ct. App. July 7, 2023) (unpublished) ("Kentucky law is clear that duty is not predicated on classifications such as 'bystander or nonconsumer or nonuser'" (quoting *Jones v. Hutchinson Mfg., Inc.*, 502 S.W.2d 66, 69-70 (Ky. 1973)).

Unable to overcome this precedent, Defendants rely on a series of inapposite cases to mischaracterize the Districts' claims as too attenuated to impose a duty. Defendants' reliance on *Gourdine v. Crews*, 955 A.2d 769 (Md. 2008); *CertainTeed Corp. v. Fletcher*, 794 S.E.2d 641 (Ga. 2016); *Quiroz v. ALCOA Inc.*, 416 P.3d 824 (Ariz. 2018); and *Kuciemba v. Victory Woodworks*, 531 P. 3d 924 (Cal. 2023) is misplaced. Each involved remote, incidental injuries to individuals unconnected to the defendants' actions. In *Gourdine*, the court declined to impose a duty on a drug manufacturer because the injured motorist was a remote bystander, with no contact or connection to the manufacturer or product. 955 A.2d at 786 ("there was no contact between [the defendant] and [the decedent] whatsoever"). Similarly, *CertainTeed*, *Quiroz*, and *Kuciemba* involved "take-home" exposure claims—arising from asbestos and the coronavirus—where the plaintiffs were physically and relationally distant from defendants' conduct. *CertainTeed*, 794 S.E.2d at 643; *Quiroz*, 416 P.3d at 827; *Kuciemba*, 531 P. 3d at 930-931.

Unlike the plaintiffs in those cases, the Districts are not remote from the conduct of Defendants. Indeed, this Court has already found "sufficient juxtaposition of the parties in time and space to place the [the Districts] in danger from the [Defendants'] acts." MTD Order, 754 F. Supp. 3d at 973–74 (citation omitted). Defendants identify no facts that undermine this Court's prior holding. Their conduct was neither passive nor attenuated: Defendants directly targeted schools and students, promoting youth engagement during school hours while knowing the predictable consequences to the Districts. *See infra* §§ A.2.a.2-3,

b.2-3, c.2-3, d.2-3 (discussing each Defendant's targeting of school-aged children and schools); Ex. 313 at 5641 (Meta: "The Young Ones are the Best Ones"); Ex. 414 at 0624 (TikTok: high schooler[s] is a core audience of our platform"); Ex. 824 (Snap "saw peaks of [Snapchat] activity during the school day"); Ex 733 at 8317 (YouTube: "Shorts is our big thing for teen appeal"). Their platforms deliberately captured minors' attention throughout the school day, making schools a direct and foreseeable locus of harm—not incidental bystanders. *See id.*; *Kennedy Krieger Inst. v. Partlow*, 191 A.3d 425, 457 (Md. 2018)(distinguishing *Gourdine* and recognizing that a duty may arise where there is a "relational link and direct contact" between the defendant and the injured party). By deliberately fostering compulsive use among students, and failing to warn the Districts, students, and parents of the risks of their platforms, Defendants directly and foreseeably caused the Districts' injuries. MTD Order, 754 F. Supp. 3d at 968, 974–76; *see also In re JUUL Labs, Inc., Mktg., Sales Practices & Prods. Liab. Litig.*, 497 F. Supp. 3d 552, 645-651 (N.D. Cal. 2020) (finding that "youth-oriented advertising" and "targeting of youth" created a foreseeable risk of harm to school Districts forced to respond to vaping on campus).

Further, this Court has already rejected Defendants' floodgates argument, finding "[t]he doomsday view is a dramatic overstatement." MTD Order, 754 F. Supp. 3d at 976. The Court held that the recognized duty is "no more expansive than defendants' own intentional and targeted actions to specific schools." *Id.* at 976–77. Thus, imposing duty here does not expand tort law. It enforces accountability for a defined category of misconduct: targeting minors through harmful platform design and failing to warn of the resulting risks. *See In re JUUL Labs*, 497 F. Supp. 3d at 658 (finding that defendants violated Arizona law and public policy "by marketing an addictive product to school-age children, leading to an epidemic of youth vaping and corresponding harms to schools"). As this Court previously recognized, "defendants' concrete and particularized awareness of potential and actual harms to [the Districts] appropriately tailors the scope of duty." MTD Order, 754 F. Supp. 3d at 976-977.

c) Defendants owe a duty to warn students and parents about their platforms' dangers.

As part of their duty to the Districts, Defendants were obligated to warn the Districts' students and their guardians of the risks posed by their platforms. Instead, Defendants deliberately fostered compulsive use of their platforms, "which foreseeably caused the plaintiff school Districts to respond by expending

resources to mitigate the impact of such use in their schools." MTD Order, 754 F. Supp. 3d at 968, 974–76; *see* Ex. 38 at 8062 (discussing Meta's goal to "saturate school networks"); Ex. 537 at 3040–41 ("[TikTok] is addictive," and "compulsive usage correlates with a slew of negative mental effects"); *id.* at 3041 ("compulsive usage on TikTok is 'rampant'"); Ex. 930 at 5370 (Snapchat's "addictiveness" affects school performance); Ex. 701 at 1255-56 (discussing YouTube's goal to "[i]ncrease habitual users" and "focus on making YouTube a daily habit").

Courts have long recognized that a defendant may be held liable for failing to warn a non-plaintiff when that failure leads to a plaintiff's injury. *See Parris v. 3M Co.*, 595 F. Supp. 3d 1288, 1337–38 (N.D. Ga. 2022) (holding that PFAS manufacturers' duty to a proposed class of landowners included a duty to warn a downstream mill of the dangers posed by PFAS); *USA Truck, Inc. v. Sullair, LLC*, 2020 WL 12656229, at *3 (M.D. Ga. Dec. 7, 2020) ("the violation of a duty to exercise reasonable care toward one person to whom that duty is clearly owed can proximately cause an injury to someone other than the person to whom the duty is directly owed"); *see also Stringer v. Nat'l Football League*, 749 F. Supp. 2d 680, 703 (S.D. Ohio 2009) ("If the manufacturer's failure to warn influenced the conduct of a third party, and that third party's acts or omissions were the proximate cause of the plaintiff's injury, then the manufacturer may be held liable.").

Even though Defendants are the movants and bear the burden on summary judgment, they present no evidence that they warned the Districts' students and their parents about the risks associated with Defendants' platforms. Nor do they contest that the Districts' injuries were a foreseeable result of this failure to warn. That alone demonstrates a genuine issue of fact for a jury. Instead, Defendants invoke three inapposite cases—*CertainTeed*, *Gourdine*, and *Quiroz*—to argue for a rule this Court has already rejected. That reliance is misplaced and contrary to established precedent as well as the law of the case.

Multiple courts—including Defendants' authority *CertainTeed* in its concurrence—have refused to adopt Defendants' broad reading. *CertainTeed* turned not on any categorical limitation on the duty to warn, but on the absence of foreseeability due to the parties' highly attenuated relationship. *USA Truck,* 2020 WL 12656229, at *3. As the court in *USA Truck* explained, "*Certainteed* did not hold that a manufacturer could not be held liable when a manufacturer breaches its duty to warn a user and that breach proximately harms a third party." 2020 WL 12656229, at *3; *see also Parris*, 595 F. Supp. 3d at 1337

(holding the reasoning in *CertainTeed* "does not foreclose a duty to warn" non-plaintiffs); *CertainTeed*, 794 S.E.2d at 646 (noting that the majority opinion "does not address whether [a plaintiff] might have a claim as an alleged injured party resulting from the failure to warn [an end user.]") (Melton, J. concurring). No such attenuation exists here.

The same is true of *Gourdine* and *Quiroz*, both of which turned on the absence of any meaningful relationship between the defendant's conduct and the plaintiff's harm. In *Gourdine*, the court found no duty where "there was no contact between [the manufacturer] and [the decedent] whatsoever." 955 A.2d at 786. Likewise, *Quiroz* emphasized that "duties are relational" and held that a company's connection to an employee's family member was too attenuated to support a duty. 416 P.3d at 839. In each case, the courts declined to impose a duty because the plaintiffs were remote from the defendants' conduct.

The factors that defeated liability in *CertainTeed, Gourdine,* and *Quiroz* are not present. As this Court has held, the relationship between Defendants and the Districts is sufficiently direct, given that Defendants' platforms were directed at students and schools; the Districts' injuries were foreseeable; and imposing a duty to warn accords with established public policy. MTD Order, 754 F. Supp. 3d at 972–78.

d) Defendants' failure to warn caused the Districts' injuries.

Contrary to Defendants' assertions, genuine issues of material fact exist as to whether Defendants' failure to warn caused the Districts' injuries. The District-specific oppositions set forth in detail the substantial record evidence demonstrating Defendants failed to warn the Districts, their students, or guardians, and the direct injuries the Districts suffered as a result. *See* Districts' MSJ Opps.; Ex. 1085; Ex. 1086; Ex. 1087; Ex. 1088; Ex. 1089.; Ex. 1090; Ex. 1083; Ex. 1106; Ex. 1107; Ex 1108; Ex. 1109; Ex. 1110; *supra* § III.A.4. As Defendants' own authority demonstrates, this is sufficient evidence of causation to proceed to trial. *See Jones by & through Jones v. IC Bus, LLC*, 626 S.W.3d 661, 682 (Ky. Ct. App. 2020) (finding directed verdict improper where there was evidence demonstrating defendant had knowledge of the dangerous condition and failed to provide any warning); *see also R & R Insulation Servs. v. Royal Indem. Co.*, 705 S.E. 2d 223, 235 (Ga. Ct. App. 2010) (finding summary judgment should be denied where there was evidence that defendants caused injury, because "[w]hile the jury may ultimately determine that any additional warning would not have prevented the harm here, that question is not for this Court to determine at this stage of the proceedings."). Unable to contend with this record, Defendants

suggest the Districts lack evidence that is not required under the law.

Defendants rely on a single inapposite case to argue expert testimony is required on the "adequacy and efficacy" of warnings here. Defs.' Breathitt Mot. at 36; Defs.' Charleston Mot. at 41; Defs.' Harford Mot. at 43-44; Defs.' Irvington Mot. at 34; Defs.' Tucson Mot. at 47 (citing *Nelson v. Am. Honda Motor Co.*, 2021 WL 2877919 (W.D. Pa. May 17, 2021)). They are wrong. In *Nelson*, the defendant had issued warnings, and the court required expert testimony to evaluate "***the contents of the existing warnings***." *Id.* at *7 (emphasis added). Other cases cited by Defendants likewise confirm that consideration about the adequacy of warnings is reserved for cases where there is an existing warning. *Allen v. Long Mfg. NC*, 505 S.E.2d 354, 360 (S.C. Ct. App. 1998) (evidence that *existing warning* was inadequate created issue of fact as to whether an adequate warning could have changed plaintiff's conduct); *Hickerson v. Yamaha Motor Corp., U.S.A.*, 2016 WL 4367141, at *5 (D.S.C. Aug. 16, 2016) (plaintiff failed to show *existing warning* inadequate where plaintiff testified she did not read it); *Georgia Cas. & Sur. Co. v. Salter's Indus. Serv.*, 734 S.E.2d 415, 421 (Ga. Ct. App. 2012) (no evidence existing warning inadequate).

Here, Defendants do not dispute that they provided no warnings about the risks their platforms pose. Because there are no warnings to evaluate, expert testimony on their "adequacy" or "efficacy" is unnecessary and would be illogical. Courts repeatedly recognize that juries are capable of understanding a failure-to-warn claim where no warning exists. *See Bullock v. Volkswagen Grp. of Am., Inc.,* 107 F. Supp. 3d 1305, 1316 (M.D. Ga. 2015) ("Plaintiffs contend that Defendants knew of risks with the design but provided no warning. A jury can understand this issue without a warnings expert."); *Sullivan v. Daimler Trucks N. Am., LLC*, 2019 WL 13563633, at *7 (D.S.C. Aug. 28, 2019) ("Plaintiff is not required to produce a human factors expert to testify about the need for a warning about a dangerous condition when no warning exists at all"); *Ruggiero v. Yamaha Motor Corp., U.S.A.*, 2017 WL 1197755, at *10 (D.N.J. Mar. 31, 2017), *aff'd*, 778 F. App'x 88 (3d Cir. 2019) (expert testimony on warnings not necessary where "contents of [existing] warnings are not at issue"); *Hanna v. Ward Manuf.*, 2016 WL 3196467, at *2 (M.D. Fla. June 9, 2016) ("[T]he Court agrees with Plaintiffs' argument that, [where no warnings were provided], expert testimony was not required to prove that Defendant's warnings were inadequate."); *see also* Fed. R. Evid. 702(a) (expert testimony must assist the jury in understanding evidence or determining a fact at issue).

Even if expert testimony were required, the record is replete with such testimony supporting the Districts' claims. Indeed, Districts' experts explain that Defendants' platforms pose unreasonable risks to children, that Defendants failed to warn *anyone* of those risks, and that this failure amplified the platforms' harmful effects, which the Districts are now forced to confront. For example, Dr. Seth Noar's expert report details the specific, foreseeable hazards Defendants failed to disclose, and explains how their omissions fell below well-established principles and standards for warnings. Ex. 997 ¶¶ 214–48. As Dr. Noar further opines, Defendants' failure to adopt and enforce a proper, standards-based warning was a driving force behind the harms children suffer from Defendants' platforms. *Id.* ¶¶ 214–64. Likewise, Tim Estes explains that Defendants could have provided clear and effective warnings about the risks of their platforms; without them schools and educators cannot effectively "deal with the fall-out from the negative mental health effects of the platforms." Ex. 989 ¶ 316. Defendants also could have, but chose not to, provide warnings in Congressional testimony and/or in materials they distributed to schools, like their guides for parents. *See infra* § III.A.2.a.6, b.6, c.6-7, d.6-7; *see, e.g.,* [redacted] at 358:14-386:6 ([redacted] testifying that CEO's congressional testimony failed to include warnings to students, parents and schools regarding the negative consequences of Defendants' platforms); Ex. [redacted] at 281:3 (same TikTok employee testifying that the TikTok Guide for Parents, which TikTok distributed to school PTAs and directly to schools, failed to warn of harms caused by Defendants' platforms).

2. Defendants breached their duty of care.

a) Meta's Conduct

(1) Meta's business model

Meta is an extraordinarily large and well-capitalized company.[3] Ex. 1 at 15-17 (over $164 billion

[3] For purposes of this and the Districts' individual opposition briefs, "Meta" refers to Meta Platforms, Inc. ("MPI") and Instagram, LLC. Meta argues that only its parent company MPI "is responsible for the development and operation" of Instagram. *See* Defs.' Breathitt Mot. at 39. But there is a genuine dispute of fact about this, precluding summary judgment. Instagram's "Terms of Use" state that Instagram is "owned or controlled by Instagram, LLC," Ex. 558 at 8784; requires consumer disputes to be directed to "Instagram, LLC," *id.* at 8788-89; and does not so much as mention MPI. Further, evidence reveals that Instagram, LLC knew about harms on the platform: Instagram, LLC contracted with a third-party research group for a quantitative study, which found that 52% of teens had negative experiences on Instagram. Exs. 1113-1115. As for the remaining named Meta subsidiaries and affiliates, the Districts and Meta are negotiating a stipulation pursuant to which they will be dismissed from the Districts' actions.

in revenue and over $5 billion in dividends and dividend equivalents in 2024). It has earned substantially all those riches from selling advertising on its platforms. Ex. 6 at 31:10-13, Ex. 7 at 338:25-339:12, Ex. 8 at 24:3-7, 248:13-249:6. However, Meta is fundamentally different than traditional forms of ad-supported media like cable TV and broadcast radio. An ad on TV is played to everyone in the same local market. Not so for Meta's platforms. Because Meta ingests an enormous amount of data about its users, it can sell advertisers the ability to "target" specific ads to specific users. Ex. 12 at 31:1-32:32; *see* Ex. 2 at 77:24-79:21 (describing types of data collected about users).

Meta has a dark history of exploiting user information to target ads. *See* Ex. 8 at 65:14-20 (former Meta advertising executive: "I helped to build an advertising system that provably, provably, could change the way that you feel about a brand of toothpaste, or any other product."). At his deposition, former Facebook research scientist (and current member of the U.K. Parliament) Dr. Joshua Simons recalled a conversation about one disturbing feature of the ad delivery system: "[T]he system knows … you're a new mum because of all the data that it has about you, and it was showing … women who just had children ads for things that would sort of make them feel bad about their body in some way, you know, like body cream or body adjustment surgery or whatever." Ex. 13 at 63:23–65:23. He explained how a colleague had worked on ways to change this but was "unable to get those things rolled out, because… they would result in fewer ads being clicked on and therefore less revenue generated by the ad delivery system." *Id.*

This abuse of user data for ad targeting is consistent with the Congressional testimony of former employee Sarah Wynn-Williams, who stated that Meta packages "sensitive personal data together and sells access to it so that advertisers can … tailor their messages to resonate with people experiencing emotional vulnerabilities or desires that an advertiser finds useful for pitching products." Ex. 202 at 3; *see* Ex. 221 at 9638-39 (Meta "marketing science" presentation examining "moments when young people need a confidence boost"). Discovery in this litigation has revealed that the source code of Meta's ranking algorithm includes the variables "Had eating disorder" and "Had weight change." Ex. 14 at 461:5-465:12. Meta's proposed algorithm expert could not rule out the possibility that Meta infers these attributes of its users and uses this information to target ads. *Id.*

Fundamentally, Meta's business model means that it makes more money the more time users spend on its apps. Ex. 8 at 93:14-94:4 ("More usage correlates with more money."); Ex. 9A at 94:5-16; Ex. 6 at

368:21-369:13. Meta is therefore incentivized to "maximiz[e] the amount of time and attention that people kind of dedicate to the platform, thus allowing for … more space and time to present the advertising." Ex. 12 at 31:8-32:10.

Testimony from multiple deponents confirms that Meta can and does attempt to increase the amount of time users spend on Instagram and Facebook to make more money. Former senior UX researcher [redacted] explained: "It seemed like when I was there, that Meta was trying to increase engagement, trying to increase how much time people spent using their apps in an effort to have more eyeballs on the apps for a longer period of time, which can get them more advertising revenue." Ex. 10 at 31:5-32:14. Dr. Simons observed that bonuses "were awarded to those teams who significantly boosted engagement," including "the amount that a user interacts on the platform, the time they spend on it, [and] the frequency … with which they visit the platform," all of which "produces more ad revenue." Ex. 13 at 134:24-13. Even Meta's since-withdrawn algorithm expert, Dr. Larry Birnbaum, conceded that Meta sets as a "goal metric" how much money it can earn from every minute someone spends on Facebook, a measure it internally calls "monetizable user value." Ex. 14 at 307:16-309:8; Ex. 15 at slide 22.

By targeting users' fundamental need for social connection—a survival instinct—and weaponizing the mathematical structure of social networks, Meta makes it nearly impossible for users to stop using their platform on their quest to monopolize their time and generate ad revenue.

(2) Meta's targeting of school-aged children

Unfortunately for a generation of young people, Meta—following the same playbook once used by Big Tobacco—determined years ago that the most valuable users, meaning those who generate the most advertising revenue, are the youngest. Ex. 313 at 5641 ("The Young Ones are the Best Ones"); *see also* Vivek H. Murthy, *Protecting Youth Mental Health: The U.S. Surgeon General's Advisory* (2023) (social media companies have "adopted tactics reminiscent of the tobacco industry"). Meta "hypothesized, that social media could be somewhat similar to, say, music where you, you know, find your tastes early on and stick with them over time." Ex. 7B at 570:9-571:16. The data bore this out. Users who join Facebook as 13-year-olds "have an LTV [lifetime value] of approximately $350, higher than all other teen ages (13-18), due to their very high long-term retention." Ex. 16 at 9025; Ex. 21 at 118:23-119:13, 176:13-178:12; Ex. 24 at 1 ("The difference in long term retention can be dramatic, with the people who signed

up for Facebook when they were young teens being 100% higher"). Teenage users presented another advantage: They could serve as gateways into the "household ecosystem," including pre-teen siblings. Ex. 18 at 1070 ("Teens strongly influenced preteens' understanding of what and how frequently to share on Instagram"); Ex. 22 at 252:18-253:9 (Instagram researcher: "teens are often the ones that other members of the household learn about not only Instagram, but social media in general through.").

Armed with these insights, Meta concluded it should "acquire teens as young as possible." Ex. 16 at 9025; *see also* Ex. 28 at 6914 (2017: "Teens: Let's always build and prioritize with them in mind."); Ex. 6 at 114:16-17, 114:22-115:3 (Instagram CEO: "Mark [Zuckerberg] is suggesting that teen time spent be our top goal in 2017."); Ex. 26 at 5014 (2018: "Winning Teens = Winning Generations."); Ex. 17 at 0071 (2021: "Youth and Teens are critically important to Instagram."); Ex. 18 at 1071 ("capturing the teen user cohort on IG is critical."); Ex. 33 at 11 (2024: "Acquiring new teen users is mission critical to the success of Instagram.").

Despite its relentless "focus[] on the teen retention and growth," Ex. 20B at 642:19-643:6, Meta's leadership knew by 2015 that the company was losing its youngest users—and treated that decline as an existential threat. Internal reports warned of a "significant DAP and engagement problem with teens," calling it "worse than we previously thought and getting worse over time." Ex. 25 at 7; *see also* Ex. 9 at 202:18-205:7; Ex. 26 at 4998 (2018: "Teen Opportunity Cost… total present value loss ~$12 billion"); Ex. 27 at 8621 ("our brand isn't faring well with teens"); Ex. 31 at 40 ("Instagram has an existential crisis with [t]eens"); Ex. 50 at 5556 ("We have a teens problem"); Ex. 20A at 241:16-242:14 ("Alarm bells started going off as youth ceased using the platform or engaging the platform as much"); Ex. 9 at 198:12-20 (addressing teen market "definitely one of Mark [Zuckerberg]'s priorities and concerns."); Ex. 6 at 126:23-127:10 (Instagram CEO Adam Mosseri: "if US teens were using Instagram less, that would be a big concern… It would be a big risk").

Rather than address the safety or well-being of these users, Meta chose to double down—knowingly sacrificing safety to recapture its most profitable audience. The company rushed to deploy new features designed to hook teens, in line with its internal mantra to "Move Fast and Break Things." Ex. 8 at 282:12–283:3 (this phrase "literally was a poster on the walls."). At Zuckerberg's direction, employees undertook a "lockdown sprint" to launch Facebook Live in early 2016 as "the beachhead we need to

expand into other use cases in video and teens," the "two spaces we care about deeply but are behind in." Ex. 32 at 9332–33; Ex. 11A at 337:8–21. Meta intentionally omitted parental and teacher safeguards, with Zuckerberg warning that notifying adults "will probably ruin the product from the start" and instructing that the company "be very good about not notifying parents/teachers." Ex. 205 at 6655. Worse still, Meta learned that Live was being used to broadcast teen suicides and attempts, an issue it did nothing to prevent or warn about before the feature's launch. Ex. 8 at 285:10-287:13. As one former Meta VP conceded, "it was a growth imperative to make Live Video a product and release it, as opposed to slowing down and really trying to think through all the negative ways it could be used." Ex. 8 at 287:15-25. Only in April 2025—nearly a decade later—did Meta add "new built-in restrictions…so that teens under 16 can't access Live without a parent's permission." Ex. 271 at 1.

Similarly, when the U.S. launch of TikTok in 2018 caused a "decline in US Teen growth"—which Meta perceived as "an existential threat for not just IG but the company overall"—it responded by rushing to release Reels, a feature that knocked off TikTok's endless feed of vertical, short-form videos. Ex. 30 at 0933; Ex. 3 at 31:7-10, 31:23-32:4 ("There were clear market competitive motivations to build that product fast"); 215:21-216:9 (describing "increased emphasis on pushing Reels" and "build[ing] this product as engagingly as possible"). The launch was a success, in the sense that millions of young users began using Reels in a matter of months. Ex. 3 at 64:11-15; *see also* Ex. 29 at 0126 ("Reels has grown and TikTok decelerated in 2022, and so we have closed a significant amount of the gap, but we are still well behind."). However, Reels was also rushed out without appropriate safety protections for young users. Ex. 2A at 38:20–24 (Q: "In your experience, was the launch of Reels on Instagram handled in a way that best protected the safety of Instagram's youngest users?" A: "No, I don't think so."); Ex. 3 at 53:11-21 ("the integrity guardrails that they put in place were not sufficient"). And Meta continued to launch changes to Reels they knew made it less safe. Ex. 206 at slide 54 ("We released a high volume (100+ in H1 2021) of product launches, 17 leading to integrity regressions."). Meta did not pause to evaluate whether those features could operate together to present even greater safety risks. Ex. 3 at 50:9-12, 52:3-53:9 (Meta "did not test for whether there were interaction effects" but "did see…later on, the cumulation of all of these launches, did lead to integrity regressions"). Meta failed to warn the Districts or the public about any of these dangers. Ex. 2A at 56:12-57:17; Ex. 3 at 86:11-19; Ex. 997 ¶¶ 142-148.

(3) Meta's targeting of schools

Meta's obsession with dominating the time and attention of youth led it to embark on a concerted, years-long campaign to embed Instagram and Facebook directly into school communities. *See* Ex. 373 at 0761 ("one of the things we need to optimize for is sneaking a look at your phone under your desk in the middle of Chemistry :)"); Ex. 37 at 4706 ("For U.S. Teens, we should focus Facebook on high school communal activities…tipping schools may be high impact"), 4714 (2016: "per-high school adoption is a crucial driver of teen Facebook engagement"); Ex. 44 at 0316 ("there's an opportunity to do classical FB growth on a high-school-by-high-school basis"); Ex. 36 at 3595 (2018: "Winning schools is the way to win with teens because an individual teen's engagement is highly correlated with school MAP [monthly active persons] penetration…Winning Teens = Winning High Schools."); Ex. 38 at 8062 (goal: "saturate school networks"); Ex. 45 at 3732 ("Over an 18 month period we expect to drive an estimated 5% teen DAP (+2.5% teen MAP) from our high school work."); Ex. 46 ("[e]ngaging the vast majority of teens in an area / school with our products is crucial to driving overall time spent in the same area"); Ex. 53 at 7145 (2023: "Can we take this a step further and use school networks as a lever for acquisition?"), 7146 (2023: "How can we position Instagram as integral to navigating school relationships, especially during transition periods?"); Ex. 55 at 5749 (2024: "we will need to invest in virality within school communities to get teens and their friends on IG."). The clear goal of Meta's campaign was to legitimize its platforms in the eyes of teachers, administrators, and parents, so that their use among students would become normalized.

Meta's school initiative took a number of different forms. ***First***, Meta developed the technical capability to determine when teen users were at school. Ex. 14 at 468:14-469:5 (Meta's expert acknowledging that this capability is embedded in the source code of Meta's ranking algorithm); Ex. 39 at 145:3-146:16 (Meta's corporate representative acknowledging that Meta is able "to look at different signals to understand when teens are online and jump to whether that is during the school day or not, or if they are physically at school"). Meta could have, but did not, use this information to reduce teens' engagement with their platforms during school hours. Ex. 14 at 470:5-471:17.

Second, Meta embarked on efforts to identify students' specific school affiliations—even though half of teens decided against sharing that information with the company. Ex. 39 at 94:10-21 ("a lot of

teens would put Hogwarts"), 97:19-4 ("only 9.3 million out of the 18.7 million active …included a school"). This has included efforts to infer students' school affiliations using complex data mapping techniques. Ex. 1179 at 7260 ("it is essential that we correctly identify the school a given teen is attending"), 7265 ("we plan to embed each teen as a data point and each school as a multidimensional gaussian" which "will allow us to directly derive the probability of attending any given school"); Ex. 43 at 2355 (describing how a teen's location can be mapped to a school, if analyzed "during school hours"); Ex. 49 at 1566 (acknowledging "FB ML [machine learning] school inference model"). It also includes a recent effort (dubbed "Project Gryffindor") through which Meta encourages teens to display their school affiliations on Instagram. Ex. 42; *see* Ex. 40 at slides 8-9 (discussing need to acquire "databases of US high schools" for this project). Meta has taken these measures because "increasing high school profile completion drives engagement and DAP [Daily Active Persons]." Ex. 41 at 4558. Indeed, Meta has estimated that "[u]nlocking new networks through school communities" has the "potential to drive ~150K DM [Direct Message] Teen DAU [Daily Active Users]." Ex. 55 at 5739-5740; *see* Ex. 36; Ex. 45 at 3731; Ex. 46.

Third, using location and school-specific information collected through data harvesting, Meta targeted push notifications to students within specific schools—which it called "school blasts." Ex. 54 at 9227. The express aim of these "blasts" was to "tip schools from inactive to active via network effects and incrementally increase teen MAP [monthly active persons] in those schools." Ex. 54 at 9227.

Fourth, Meta paid the National PTA and Scholastic to conduct an extensive outreach campaign targeting schools and families. Meta chose these partners precisely because they were "trusted, respected organizations" with "significant credibility" that could "be a public validator" for the company and "get our materials into the hands of parents, grandparents, and educators at scale." Ex. 48 at 4519. Meta's agreements with these organizations "include significant website placement, social media promotion, targeted physical distribution, community activations, and post-campaign surveys." *Id.*; Ex. 47 at 4527. Meta's conduct mirrors the tactics of Big Tobacco, which used "youth smoking prevention" and school-based programs to infiltrate classrooms and legitimize its products under the guise of education—a strategy condemned in *United States v. Philip Morris USA Inc.*, 449 F. Supp. 2d 1, 853–71 (D.D.C. 2006)

(finding tobacco companies intentionally marketed to youth through schools while falsely claiming to promote public health).

Meta "began a relationship with the National PTA as early as 2014, possibly earlier." Ex. 39 at 25:6-26:3. In 2018 alone, Meta paid the PTA $110,000, with part of that budget "coming from marketing." Ex. 48 at 4519. In exchange, Instagram was made an "Official partner for PTA's first-ever Back to School Week campaign," given a "Dedicated Back to School landing page housing Instagram-supplied educational content," and benefitted from an "Email blast[s]" to "~16K PTA presidents at middle and high schools nationwide" and "~150k PTA members." Ex. 48 at 4520. Internally, Meta praised the partnership for granting access to the PTA's "16.2 million families across the U.S.," noting that "If you want to connect with parents at scale, you work with the PTA." *Id.* Through its PTA partnership, Instagram distributed parents' guides, hosted focus groups with educators and administrators, and held events at "nearly 200 schools across the country, in every single state." Ex. 39 at 190:11-192:20; *see also* Ex. 51 at 7467 (planning conversation with PTA on "specific sponsorship amount" and "parent roundtables"), Ex. 52 at 6709 (discussing "Facebook PTA Digital Toolkit Campaign"). Meta even worked with the PTA to "[c]apture footage of PTA families trying out" new Facebook products "and offering testimonials"— to "build trust with parents and with press." Ex. 277 at 5130.

Meta's paid relationship with Scholastic mirrored its partnership with the National PTA in both scale and purpose. Ex. 56 at 6341 ("we worked with Scholastic in a big way for our last IG parents' push"). In 2018, Instagram's marketing team paid Scholastic $135,000 and, in exchange, received "Physical distribution" of its materials "to 20,000 classrooms across the U.S. (Reach: 20,000 teachers / 600,000 students/ families assuming 30 students per class)." Ex. 48 at 4519. Meta recognized that Scholastic's reputation would amplify its message, noting that "98% of teachers feel confident about sharing information that they receive from Scholastic with parents of their students." Ex. 48 at 4520. As Meta explained internally, "Scholastic gives us the ability to win educators and the school community, creating a ripple effect to allow us to win parents and families. Educators and schools don't feel like they have enough information to help teens in their schools; a Scholastic partnership helps fix that." *Id.*; *see also* Ex. 50 at 5583 ("Leverage established relationship with Scholastic to facilitate activations").

Fifth, Meta engaged in a deliberate campaign of direct outreach to schools and school districts to normalize its products under the guise of education and safety—again mirroring the tactics once used by Big Tobacco to gain access to youth audiences. The company presented "safety roadshows" at high schools across the country, including Evergreen High School and Miami-Dade County Public Schools. Ex. 57 at 3853; Ex. 58 at 3188. It hosted "Teacher Institutes" for 400 educators at its offices (Ex. 64 at 8612); addressed "THE conference for charter schools around the US." in partnership with the National Alliance of Charter Schools (Ex. 65 at 0703; Ex. 59 at 4307); and created a "digital citizenship" program "[d]esigned with teachers in mind," piloted in multiple middle schools (Ex. 60 at slides 3, 7; Ex. 61). Meta launched a Facebook-powered "personalized learning platform" in select charter schools, through Summit Public Schools. Ex. 62 at 4392. It piloted an experimental teen-only app in "the 10 largest schools in NYC and Miami." Ex. 63 at 5481. Even after this lawsuit was initiated, its efforts have continued; in 2024, Instagram launched what it calls a "School Partnership Program." Ex. 39 at 39:3-10, 40:12-14, 18-21; *see also* Ex. 73 at 1695 (Instagram head of policy noting "a lot of conversations lately with Principals at schools and school districts").

Sixth, Meta undertook marketing efforts specifically designed to reach teenagers in their role as students. This included a "Back-to-School with IG" brand campaign that "showcase[d] teens navigating the transition between summer vacation and back-to-school through product-centric storytelling." Ex. 66 at 2285; *see also* Ex. 67 at 6488 ("launching a new Back-to-School hub"). Meta's outreach went further: the company recruited and paid 13–17-year-old "teen tastemakers to act as our plug at local high schools within 5 key markets." Ex. 68 at 6448; *see* Ex. 71 at 0328 ("US Teen Lifetime Value" presentation asking, "What's a reasonable level of funding for a teen ambassador program?"). These "ambassadors" received compensation—including Amazon gift cards, branded "swag" like hats and sticker sheets—and were required to sign nondisclosure agreements. Ex. 68 at 6453-54, 6449. Meta used these teen ambassadors to "build excitement at local high schools" for products such as its "high school directory" pilot (Ex. 69 at 0219, 0221, 0241) and Reels (Ex. 70 at 9347).

As the foregoing should make clear, Meta had every conceivable opportunity to put school administrators, teachers, parents, and students on notice if it knew about safety and health risks associated with students' use of Instagram and Facebook. It had every conceivable opportunity to warn schools if it

knew those risks could materialize in the classroom. Meta did not do so. Instead, it used paid partnerships with validators and generic educational toolkits to infiltrate school communities, legitimize Facebook and Instagram, and promote platform adoption (including by schools themselves)—with the predicted and desired result of growing teen engagement to drive long-term revenue.

(4) Meta's knowledge of harms

What Meta knew—and what it failed to reveal—should shock the conscience. Internal studies conducted by Meta, both qualitative and quantitative, repeatedly confirmed that Instagram and Facebook use could lead to a panoply of negative outcomes for teens, including addiction, sleep disruption, anxiety, depression, negative appearance comparison, and body image problems.

We begin with addiction. And we start with what is perhaps the most succinct acknowledgment of the problem, from a chat between two UX Researchers: "oh my gosh yall IG is a drug…We're basically pushers… We are causing Reward Deficit Disorder bc people are binging on IG so much they can't feel reward anymore…like their reward tolerance is so high…I know Adam [Mosseri] doesn't want to hear it – he freaked out when I talked about dopamine in my teen fundamentals leads review but its undeniable! Its biological and psychological….the top down directives drive it all towards making sure people keep coming back for more. That would be fine if its productive but most of the time it isn't…the majority is just mindless scrolling and ads." Ex. 74 at 2152-53.

Far from being one-off musings, these insights are echoed in a chorus of evidence from other Meta researchers. *See* Ex. 75 at 1420 ("Teens talk of Instagram in terms of an 'addicts narrative' spending too much time indulging in a compulsive behaviour that they know is negative but feel powerless to resist."), 1428 ("Teens are hooked despite how it makes them feel. Instagram is addictive, and time-spend on platform is having a negative impact on mental health."); Ex. 76 at 3172 ("56% of IG teens surveyed say it's difficult to manage how much time they spend on social media and 14% say IG makes it worse"); Ex. 85 at 5016 ("Teens may sometimes be using social media more than they intended to without paying attention to it, such as engaging in habitual use....Teens may also believe that it's challenging to limit their use or feel like it's hard to exert self-control"); Ex. 19 at 77:24-78:6, 80:15-23, 128:19-129:3 ("It wasn't that teens didn't like it. It's that they didn't think that it was good for them. The teens that we talked to … thought that they should be controlling it. They didn't think that they could."); Ex. 78 at 7036 ("(1) teens

feel addicted to IG and feel a pressure to be present, (2) like addicts, they feel that they are unable to stop themselves from being on IG, and (3) the tools we currently have aren't effective at limiting their time on the ap[p]"); Ex. 13 at 67:2-10 ("addiction was one of the types of harm…that I talked about with senior executives"); Ex. 77 at 3673 ("sophisticated reinforcement learning might increase the problem of tech addiction in Stories, News Feed, or Instagram"); Ex. 79 at 4205-4206 ("Are there teens that use too much Instagram? It sure looks like it….It seems relatively clear that younger people are less equipped to handle social media addictions….In the US alone, there are 550K+ teens that spend 28 hours a week on Instagram (4 hours a day, 7 days)."); Ex. 110 at 3026 ("it is clear that clinicians are seeing people reporting problems with Facebook use."); Ex. 132 at slide 32 ("social media addiction: heavy usage of Facebook correlates with lower happiness").

Meta's understanding of the addiction issue dates at least back to the summer of 2017, when a data science intern internally published a research note titled, "Have we made people addicted to Facebook?" Ex. 104 at 0086; *see* Ex. 140. As part of this research, the intern and his supervisors spoke to the author of a book called *Hooked* to "discuss[] his extensive research into creating habit-forming products[.]" *Id.* The team then fielded a survey, received responses from 1,300 Facebook users, built a model, and determined that "we could call addicted…nearly 200k people." Ex. 104 at 0089. The intern emailed Adam Mosseri—then the Head of News Feed at Facebook, now the CEO of Instagram—asking to discuss his findings. Mr. Mosseri blew him off. Ex. 6 at 391:15-23 ("can't meet with every intern that reaches out.").

Just weeks later, two of Meta's former senior executives offered scathing critiques of Facebook's addictiveness. On November 9, 2017, Sean Parker, "the second president" of Facebook, Ex. 11A at 48:3-14, made public statements that "[t]he thought process that went into" building Facebook "was all about how we consume as much of your time and conscious attention as possible. And that means that we need to sort of give you a little dopamine hit every once in a while because someone liked or commented on a photo or post or whatever. And that's going to get you to contribute more content…it's a social validation feedback loop…it's me, it's Mark…[we] understood this consciously, and we did it anyway." Ex. 134. "God only knows what it's doing to our children's brains." Ex. 135. At his deposition, Mark Zuckerberg credited Mr. Parker as "a smart person" who "helped out a lot with the business" and who influenced his thinking around "business strategy." Ex. 11A at 48:15-49:18, 52:3-8; *id.* at 51:10-18 & Ex. 136

("massively influential on my thinking"). Nonetheless—even though Mr. Zuckerberg "saw the headlines" around Mr. Parker's comments—he did not reach out to him to discuss or "look into anything based on this." *Id.* at 55:2-19, 55:25-56:6 ("I don't need to investigate that.").

On November 13, 2017, another Facebook executive went public with his concerns—Chamath Palihapitiya, a vice president at Facebook from 2007 to 2011 with whom Mr. Zuckerberg was in regular communication during his tenure. Ex. 11A at 93:24-94:15. Mr. Palihapitiya stated that "The short-term, dopamine-driven feedback loops that we have created are destroying how society works" and "eroding the core foundations of how people behave." Ex. 137. "We curate our lives around this perceived sense of perfection because we get rewarded in these short-term signals, hearts, likes, thumbs up, and we conflate that with value….And, instead, what it really is, is fake, brittle popularity that's short-term and that leaves you even more…vacant and empty before you did it. Because then it forces you into this vicious cycle where you're, like, what's the next thing I need to do now, because I need it back." Ex. 137. After the press reported on his comments, Mr. Palihapitiya wrote to Mr. Zuckerberg apologizing for "all the pickup today" (but conspicuously not disavowing the substance of his remarks). Ex. 138 at 6409-10. Mr. Zuckerberg did not respond to the email and has never spoken to Mr. Palihapitiya since. Ex. 11A at 103:23-104:7, 106:12-17 & Ex. 1181.

While Meta's leadership ignored growing evidence of a problem, researchers at the company began to undertake a series of qualitative, quantitative, longitudinal, and even experimental studies confirming the widespread phenomenon of what it euphemistically called "problematic use." *See* Ex. 78 at 7036-37 ("Do we want to call it addiction?" "we would never want to say that!"); *but see* Ex. 107 at 0181 (acknowledging that problematic use "is sometimes referred to as 'social media addiction' externally."). Meta internally defined problematic use as "experiencing both of the following issues either 'very often' or 'all the time': • Lack of control or feelings of guilt over Facebook use. • Negative impact in at least one of the following areas: sleep, parenting, social relationships, or productivity." Ex. 80 at 5381. "Behaviors associated with PU [problematic use]" include "Frequent checking," "Overall amount of time spent," and "Passive late night use." Ex. 76 at 3173; *see also* Ex. 109 at slide 16-17 (several behaviors (e.g. total time spend, late night use, video chaining) and surfaces (e.g. Explore, Reels, notifications) "associated with low perceived control on IG.").

In 2018, Meta's researchers "paired a survey of 20,000 U.S. Facebook users measuring perceptions of problematic use with demographic and behavioral data for the prior month." Ex. 83 at 0059. That study—whose results were shared directly with Mark Zuckerberg and Meta COO Sheryl Sandberg—provided a "deep understanding" that the prevalence of problematic use among U.S. Facebook users is "55% mild, 3.1% severe." Ex. 81 at 0761 (email to Mark Zuckerberg), Ex. 82 (attachment); *see* Ex. 11A at 274:24-275:6 (Zuckerberg: "3 percent of billions of people is a lot of people…it's millions of people."). Meta never publicly disclosed the true results of this study. It did, however, publish research falsely stating, "we estimate (as an upper bound) that 3.1% of Facebook users in the US experience problematic use." Ex. 97 at 2 (emphasis added); *see* Ex. 11B at 680:9-683:8.

In 2019, Meta's researchers "extended our measurement framework" and "revised our survey questions" to "follow the same format … as is standard practice in mental health assessment scales." Ex. 99 at 0168. Using this framework, assessed against a "large, multi-national dataset" of 42,000 participants, Meta determined that "problematic use appears to affect 12.5% of people" and "[a]nother 11% of people report negative life impact 'very often' but [report] only 'sometimes' lacking control." Ex. 99 at 0169-70; *see* Ex. 80 at 5381 ("12.5% of people worldwide experience problematic use"). Meta has never publicly disclosed the results from this study.

At the end of 2019, a lead UX researcher at Facebook completed a 24-person qualitative study, consisting of "1:1 interviews with 'problematic' Facebook users." Ex. 107 at 0178. Participants described their addiction in unsparing terms: "The time I spend is not healthy, it's like an addiction. Opening the app every half hour, it's not healthy." Ex. 107 at 0196. "When you have a beer on the weekend, you drink it. You have to. It's the same thing with Facebook." *Id.* "[Y]ou can get into this rabbit hole and might not get out." *Id.* "I spend a lot of time on the videos because they start automatically." Ex. 107 at 0201. "People liking things can be addictive. I feel compelled to see who liked it. I think it's a bad habit because [I'm] always checking." *Id.* "I'm on Facebook everyday, every moment. Literally, every moment; just not when I'm in the shower. It's a habit…I lose the notion of time." Ex. 107 at 0202. While some participants attempted "fruitless" self-remediation tactics to alleviate these problems, researchers reported that "none had found a full solution" and most resorted to "deleting the [Facebook] app or deactivating at least once to curb use." Ex. 107 at 0204. Meta did not disclose these findings to the public.

Around the same time, a senior UX research leader at Instagram (and a PhD in social psychology, Ex. 11A at 83:6-19) prepared a work plan that included a recommendation that Meta warn the public about its research findings: "Because our product exploits weaknesses in the human psychology to promote product engagement and time spent [and that because of our reliance on AI, we cannot currently control for when content is being strategically leveraged to manipulate the opinions and moods of individual people] if we are committed to the wellbeing of individuals, we need to a) alert people to the effect that the product has on their brain (for example: your view of yourself is affected by other people's highlight reel, etc) and b) provide them with tools to have more control over their own experience." Ex. 105 at 4015 (brackets in original). Meta did not follow this recommendation and did not disclose the known effects of its products on the brain. Ex. 11A at 87:2-93:1.

In addition to studying problematic use, Meta conducted parallel analyses on user well-being. In 2019, the company surveyed 30,000 Instagram users—once in May and again in October—to measure the reach and intensity of perceived negative experiences on the platform. Ex. 100 (May 2019 results); Ex. 101 (Oct. 2019 results); Ex. 102 at 36:25-37:10 ("We were aiming to measure subjectively bad experiences."). The findings were alarming: "the perceived reach of negative experiences on Instagram is quite high," Ex. 102 at 44:23-45:14; Ex. 101 at slide 6, and "teens have significantly more negative experiences than adults … for 11 out of the 18 problem areas that we asked about," Ex. 101 at 7.

Meta also conducted an analysis of 13- to 17-year-olds that "was meant to be a cross section of Instagram users" and "[a] representative sample to the best of our ability." Ex. 19 at 103:10-13, 104:4-15; Ex. 98 at 9175 (methodology included survey of 2,503 users). The results were stark: "One in five teens say that Instagram makes them feel worse about themselves." Ex. 98 at 9193. Of those surveyed, "39% of teens who had felt that they needed to create a perfect image said that that feeling started on Instagram," Ex. 19 at 106:20-24; "41% of teens who felt not attractive said that that feeling started on Instagram," *id.* at 106:25-107:3; "24% of teens who felt that they were not good enough said that that feeling started on Instagram," *id.* at 107:13-16; "21% of teens who said they felt alone or lonely said that that started on Instagram," *id.* at 107:25-108:3; "10% of teens who said they were down, sad or depressed said that that feeling start[ed] on Instagram," *id*. at 108:4-7; "6% of teens who said that they felt that they wanted to kill

themselves said that that feeling started on Instagram," *id.* at 108:11-14; and "9% of teens who said that they felt that they wanted to hurt themselves said that it started on Instagram," *id.* at 108:15-18.

Other internal studies conducted during the same period confirmed these patterns. Roughly one-third of teens who reported problematic social media use said that Instagram made that experience worse. Ex. 102 at 151:14–152:6; Ex. 123 at slide 19. In another survey of 1,000 teens, 23% said they "waste too much time" on Instagram, and 19% said they "miss out on life" often or very often because of it. Ex. 111 at 32; Ex. 19 at 58:6–25. Taken together, these studies left Meta with unmistakable evidence that its platforms were worsening teen well-being.

Meta did not voluntarily disclose the results of these 2019 qualitative surveys. The findings became public only after a whistleblower leaked them to Congress and the press in 2021, prompting widespread outrage. Ex. 23 at 125:11–126:6; Ex. 209 at 9310. In response, Meta released "annotated" versions of the presentations designed to downplay and obscure their conclusions. Privately, the authors of this research (Meta's own employees) were unsparing in their assessment of this move. *See* Ex. 102 at 209:23-300:4 (author of May and October 2019 presentations: "they're absolutely throwing us under the bus and making us sound incompetent"), 301:5-8 ("they're basically trying to discredit the work so much that it's like it didn't get leaked at all."); Ex. 1182 at 1964 ("I understand why they have to defend against the causal claim externally, but internally why is it so hard to take responsibility for shitty experiences that users have on the platform that we absolutely could do more to limit? …I think their perception is that the worst thing here is the leak itself, followed by the language/mischaracterization of the article, followed by the actual problem facing humans and the fact that we haven't devoted enough resources to it."); Ex. 19 at 307:6-21 (author of teen survey work: "it makes it sound like I was just guessing"), 308:12-15 ("that phrasing of maybe researcher speculation. Like, that is the part that sat very unwell with me then and kind of still does now."); Ex. 103 at 5013 (co-author of research: "like PLEASE just focus on what they are REALLY criticizing you about and acknowledge it and stop throwing your researchers under the bus").

Meta's research into user harm extended beyond qualitative studies. In late 2019, it "designed an experimental deactivation study, in which we will randomly ask some people to stop using Facebook and Instagram for a month (compared to a group who will continue to use as normal), helping us explore the impact that our apps have on polarization, news consumption, well-being, and daily social interactions.

We'll partner with Nielsen and use a combination of (a) surveys, (b) log data from Facebook and Instagram, and (c) usage data from smartphones." Ex. 87 at 5647; Ex. 88 at slides 2-5 (study design); Ex. 89 (Nielsen contract); Ex. 92 at 2432 (codename "Project Mercury"). Meta undertook this in-house experimental study to fix what it perceived as shortcomings in the external literature. Ex. 96 at 0466 ("our design is of much higher quality"); Ex. 87 at 5647 ("one of our first causal approaches to understand the impact that Facebook has on people's lives….Everyone involved in the project has a PhD").

To Meta's disappointment, pilot tests of the deactivation study confirmed that "[p]eople who stopped using Facebook for a week reported lower feelings of depression, anxiety, loneliness, and social comparison." Ex. 90 at slides 2, 6-7. Rather than investigate further, or sound the alarm, Meta halted the project—claiming that participants' feedback was biased by "the result of the existing media narrative around the company." Ex. 93 at 0429; Ex. 94 at 3905. *But see* Ex. 93 at 0429 (Nick Clegg: "is it really possible to claim that the effect is *all* due to media conditioning/perception?"); Ex. 95 at 1927 ("the nielsen study does show causal impact on social comparison 😳").

One Meta employee warned, "if the results are bad and we don't publish and they leak, is it going to look like tobacco companies doing research and knowing cigs were bad and then keeping that info to themselves? I went…oh." Ex. 91 at 1943. That is exactly what Meta did. The company never publicly disclosed the results of its deactivation study. Instead, Meta lied to Congress about what it knew. *Compare* Ex. 105 at 4007 (2017: listing "Increase in anxiety" among the "Known Negative Effects of FB and/or Social Media in General on Teens"); Ex. 98 at 9196 (2019: "young people openly attribute their increased level of anxiety and depression to Instagram"); Ex. 90 at slide 2 (Sept. 2020: cessation of Facebook use leads to lower depression and anxiety) *with* Ex. 106 at 124-125 (Dec. 2020 response to questions from U.S. Senate: Q: "Is Facebook able to determine whether increased use of its platform among teenage girls has any correlation with increased signs of depression within this demographic…?" A: "No." Q: "Is Facebook able to determine whether increased use of its platform among teenage girls has any correlation with increased signs of anxiety within this demographic…?" A: "No."); Ex. 11 at 238:20-242:19.

Meta's internal research, and its external deception, continued. In December 2020, a member of Meta's core data science team (who holds a PhD in economics, Ex. 141 at 2) authored an internal study acknowledging "[t]here are reasons to worry about self-control and use of our products" and presenting a

"quick rundown of evidence"—including "[a]n experiment [which] found that a 1-month break from Facebook improved self-reported wellbeing." Ex. 108 at 5595 & Ex. 139 at 0119. In response, another senior data scientist at Meta (who also holds a PhD in neuroscience, Ex. 142 at 2, and taught a university course on addiction, Ex. 143 at 1) warned: "It seems clear from what's presented here that some of our users are addicted to our products. And I worry that driving sessions incentivizes us to make our product more addictive, without providing much more value. How to keep someone returning over and over to the same behavior each day? Intermittent rewards are most effective (think slot machines) reinforcing behaviors that become especially hard to extinguish—even when they provide little reward, or cease providing reward at all." Ex. 139 at 0127; *see also* Ex. 13 at 67:12-17.

Meta never publicly acknowledged these findings from its own in-house addiction specialists. Instead—just weeks later—Mark Zuckerberg testified before Congress and denied that Meta profits from creating addictive products. Ex. 144; Ex. 11A at 154:16–155:11. This deliberate contradiction between Meta's internal science and its public testimony exemplifies a broader pattern: the company knew its platforms fostered addiction and psychological harm but chose to deny, obscure, and mislead.

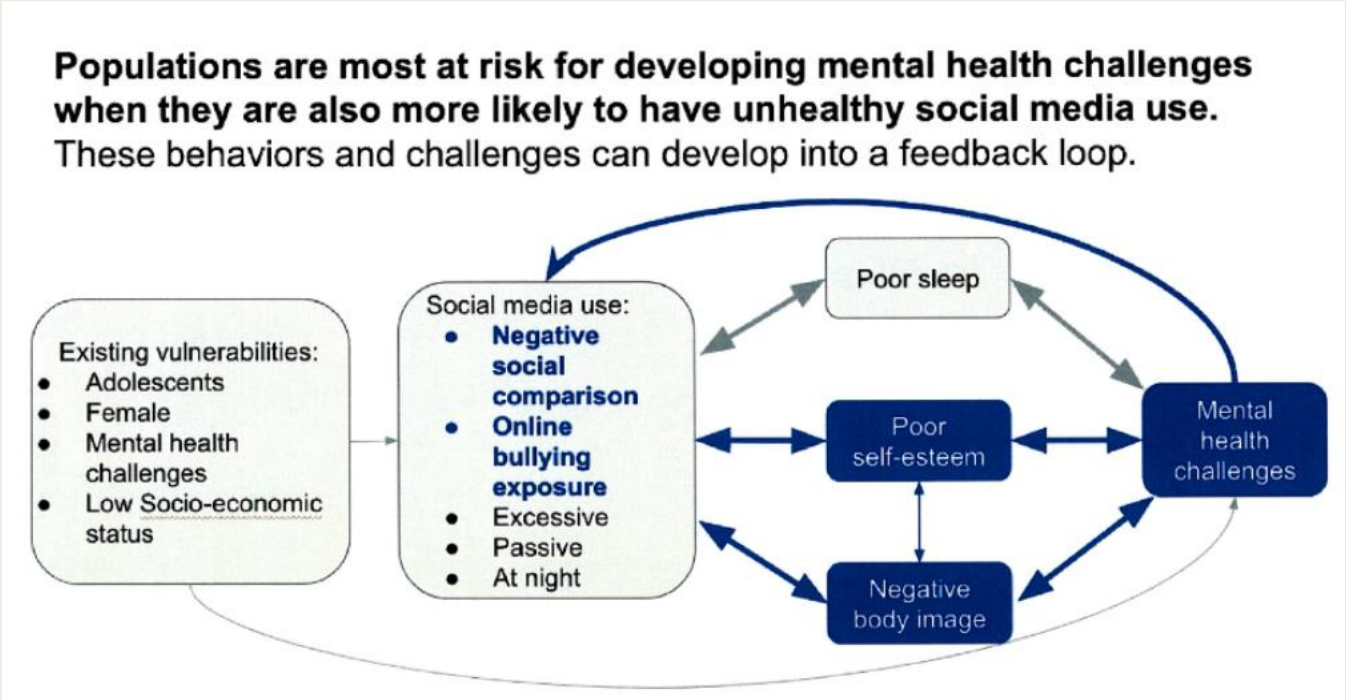
In 2022, Meta conducted a "mixed method study," surveying and interviewing over 1,000 "mental health clinicians (including psychiatrists, psychologists, therapists, social workers)." Ex. 128 at slides 5-6. Eligible clinicians had "2+ years experience post-licensure" and "Provided care for at least 30 patients in the past 3 months." *Id.* This study revealed that "the majority of clinicians believe that social media can be addictive," with fully 85% of U.S. clinicians endorsing this proposition. *Id.* at slide 9.

In 2023—"following resolution of a senior leadership-level legal escalation"—Meta's researchers launched a longitudinal study to follow 4,000 U.S. teenagers over the academic year. Ex. 84 at 7080 ("Meta & Youth Social Emotional Trends" study); *see also* Ex. 86 ("I really prefer not to have a cute acronym (MYST). That makes things sound more interesting when/if it leaks."). The goal of this survey was to measure teens' *attentiveness* ("Possible unintended, unaware, and/or habitual use") and *capability* ("Perceived self-efficacy to exert control, if desired"). Ex. 85 at 5018-19. Early results were troubling: more than "one-third (37%)" of teens reported low attentiveness on Instagram, while slightly less than "1 in 3 teens (28%) reported experiencing challenges with capability." Ex. 85 at 5015, 5019. The research also confirmed that parents and family circumstances are not to blame for this problem. Ex. 85 at 5021

("Parental and household factors have little association with teens' reported levels of attentiveness to their social media use….family income, parental marital status, and other household characteristics…were not associated with the teens' feeling of attentiveness."), 5026 ("There is no association between either parental reports or teen reports of parental digital caregiving/supervision and teens' survey measures of attentiveness or capability."). Meta has never disclosed the results from this study.

That same year, other researchers conducted a 12-country survey of 30,000 adults and 30,000 teens (13-17) on Facebook and Instagram, to identify "on-platform behaviors associated with perceived low control" over platform usage. Ex. 109 at slide 2, 10. Consistent with the MYST study, nearly 40% of youth reported "low attentiveness" and 20% of youth reported "low capability" on Instagram. Ex. 109 at slide 13; Ex. 11A at 227:11-21; *see also* Ex. 982 ¶ 372 ("Three to six hours of social media at ages 13-16 was associated with a 60% increased relative risk of internalizing problems at ages 14-17 and greater than 6 hours per day was associated with a 78% increased risk."). This research went on to assess how such compulsive use interfered with daily life—specifically, whether it prevented individuals "from fulfilling [their] responsibilities in other areas of [their] life (for example, work, school, or relationships)." Ex. 109 at 9. The results were stark: users with "low attentiveness" reported 20–25% higher life interference, and those with "low capability" reported 17–23% higher interference. Ex. 109 at slide 13. Meta has never disclosed these results.

As the foregoing study demonstrates, Meta understands not only that its platforms are addictive to teens, but also that this addiction produces a cascade of related mental health harms. In 2020, an Instagram researcher (who holds a doctorate in public health, Ex. 113A at 328:10-329:6) developed a "conceptual model of adolescent social media behaviors and mental health," in collaboration with his colleagues. Ex. 112 at 2-3. The model shows that problematic social media use is linked to poor sleep, low self-esteem, negative body image, and mental health challenges—each of which fuels further problematic use, creating a self-reinforcing cycle. Ex. 112 at 3; Ex. 113A at 341:20–343:21 ("there's a feedback loop with mental health challenges"). It also identifies adolescents with "existing vulnerabilities" as particularly at risk, whether those vulnerabilities operate independently or in combination. Ex. 113A at 337:1-15.



Ex. 112 at 3.

To be clear, the connection between problematic use and mental health issues for teenage users was not a one-off observation by a single researcher. Far from it. For example, Meta knew that addictive use of Instagram and Facebook could cause disrupted sleep and poorer mental health for young users. Ex. 114 at 9142 ("Nighttime social media use is associated with poorer mental health in teens due to displacing sleep; this is the most direct relationship between social media use and teen well-being. The most vulnerable teens offline are also the most vulnerable teens online"); Ex. 115 at 2192 ("For late night use, the negative impacts, such as on work/school performance and mood, or little or poor sleep (especially on young people) are well-documented."); Ex. 116 at 10 (43.3% of teens have at least one session on Instagram between 12am-4am each week and 4.1% of teens "spend 18+ min on IG" during these hours), 15 (two thirds of teens who exhibit problematic late-night use "start a session from a notification"); Ex. 117 at 8 (2018 feedback from "US teen": "When I'm sleeping, I get a lot of Instagram notifications"); Ex. 118 at 9909 ("As session count increases, more and more people report feeling like they spend a lot of time on Facebook or report sleep problems."). Despite this extensive body of evidence, Meta remained silent, failing to warn the Districts and the public. Ex. 997 ¶¶ 142–148.

As discussed above, Meta also knew that addictive use of its platforms contributed to depression and anxiety. Ex. 90 at slide 2; Ex. 105 at 4007; Ex. 98 at 9196. Internal research captured this vividly: teens described "The Me on Instagram" as "one-dimensional, boxed in (small box), not in control, dark, alone, pulled down, low esteem, rabbit hole, anxious, overthinking." Ex. 119 at 1117; *see also* Ex. 19 at 85:7-12 ("some teens have told us that they feel anxiety around their Instagram usage"). Meta's 2022 survey of 1,000 mental health clinicians confirmed what its own researchers already knew—81% of clinicians said social media exacerbated patients' anxiety disorders, and 78% said it worsened depressive disorders. *Id.* at slide 24. Despite this extensive evidence, Meta concealed the truth from the public and lied to Congress about what it knew. Ex. 106 at 124-125.

Over years of internal research, Meta repeatedly found that Instagram use drove negative social comparison and eroded teen self-esteem. Ex. 81 at 0761 (email to Mark Zuckerberg: "deep understanding" of "negative drivers that occur frequently on FB" including "Social comparison (prevalence: 40% mild, 5% severe)"); Ex. 127 at 8 (internal survey of 5,793 randomly selected participants in 2018 revealed that "33% of people have been feeling worse about themselves on IG for 'several months to a year'), 24 ("regression…shows that age and gender are the strongest predictors of negative social comparison"); Ex. 111 at 32 (32% of surveyed teens felt "like I have to look perfect" on Instagram); Ex. 19 at 58:6-25; Ex. 119 at 1098, 1101 ("Social comparison is worse on Instagram"; "66% of teen girls on IG experience negative social comparison"; "32% of teen girls said that when they felt bad about their bodies, Instagram made them feel worse"); Ex. 120 at 8700 (internal survey of over 50,000 users finding that "One-quarter of people on IG (23%) feel a lot or extreme pressure to look perfect on Instagram, and one-third of teen girls (34%) do," and that "Nearly half of teen girls (48%) often or always compare their appearances on IG"), 8704 ("Appearance-based comparison on Instagram is common and can have significant implications for individuals"); Ex. 102 at 153:16-24 ("somewhere between, you know, 15 and 20%" of respondents said that Instagram made social comparison worse); Ex. 128 at slide 42 (mixed-methods study of 1,000 clinicians: "Encouraging negative comparison was most cited as negative impact on mental health patients"); Ex. 129 at 1890, 1895 ("Results from a 100k-person survey + behavioral data" finding that "About 1 out of 10 people experience negative social comparison on Instagram *often or always*") & Ex. 130 at 217:22-218:2 (this is millions of people). Despite this extensive body of evidence, Meta never

warned the public. Ex. 997 ¶¶ 142–148. Instead, it continued to promote Instagram as a positive and empowering space for teens—knowing full well that its own research showed the opposite.

Nor did Meta disclose its knowledge of a related harm: that addictive use of Instagram fuels body image issues for teens. Ex. 121 at 5133 ("research suggests a causal link between Instagram use and appearance comparison and/or body image issues."); Ex. 131 at 6265 ("external research consistently points to appearance comparison as a key step towards internalization of the 'thin ideal' and other beauty ideals which often lead to body image issues and for some eating disorders/ suicide/ self injury."); Ex. 123 at 5021 (survey of 22,410 users revealing that "we make body image issues worse for 1 in 3 teen girls" and that "[m]ore teen girls thought that IG made body image issues worse") & Ex. 122A at 204:5-18, 214:21-215:2; Ex. 124 at 1221 ("the team's vision should be to move IG out of the position of being uniquely worst for teens' social comparison and body image") & Ex. 20B at 502:5-503:21; Ex. 125 at 0058 ("it seems patently obvious that phone usage, selfie-culture, and photo-editing are contributing to anxiety, depression, and body dysmorphia….we have to assume we are contributors / amplifiers."); Ex. 126 at 7066 ("there is substantial evidence to suggest that Instagram and Facebook use can increase body dissatisfaction."). Meta did not warn the public, including the Districts, about these problems.

Finally, Meta knew that addictive use of its platforms could lead to self-injury, including eating disorders and suicidality. Ex. 173 at slide 7 (discussing how "addiction" on its platforms "can lead to more downstream effects like self-injury or suicide" and "eating disorders"); Ex. 132 at slide 30 ("model for how social media usage may be linked to suicide/self-harm"); Ex. 128 at slide 24 (showing that over 30% of clinicians believed that social media "had a negative role" in suicidal behavior disorder and non-suicidal self-injury disorder). This is particularly concerning since Meta also knew that teens represented the "vast majority of users who admit and promote self-harm and suicide are teens (45%-60%)" and "reporters of self-harm and suicide are also disproportionately represented by teens (52%-56%)" despite "only mak[ing] up 18% of DAP." Ex. 172 at 0010. Meta did not warn anyone about these problems.

In short, Meta understood that use of Instagram and Facebook could lead to serious, negative mental health harms—particularly for teenage users. Ex. 114 at 9141 ("best research suggests that Facebook is net bad for well-being"); Ex. 108 at 5595 & Ex. 133 at 71:2-72:15 (internal study stating, "Higher Facebook use is correlated with worse psychological states (wellbeing, loneliness, etc.)"); Ex.

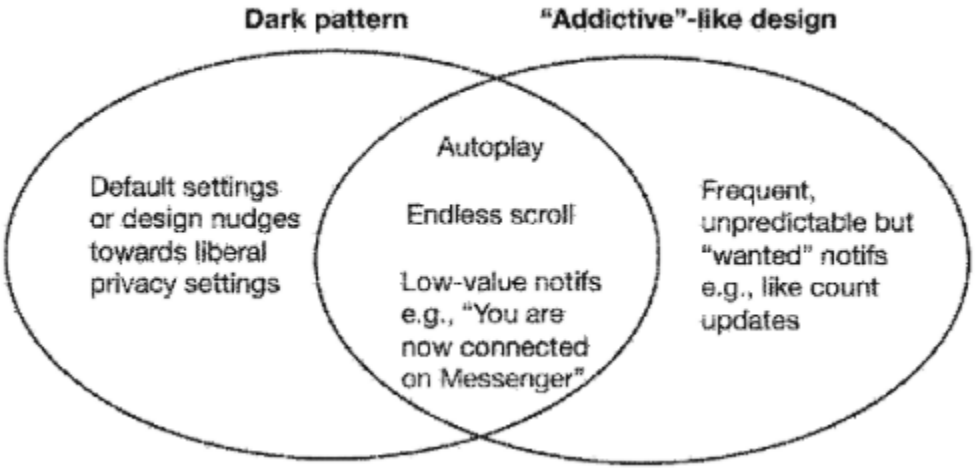
119 at 1124 ("Users experience of downward spiral is exacerbated by our platform"), 1125 ("Aspects of Instagram exacerbate each other to create a perfect storm"), 1126 ("Mental health outcomes related to this can be severe" and include "depression," "body dysmorphia," and "eating disorders"); Ex. 19 at 165:10-166:9 ("teen girls that we talked to very clearly associated this spiral that they were experiencing on Instagram with some, you know, pretty real – pretty bad outcomes."), 167:19-25 ("the teen girls and young adult women that we talked to felt as if Instagram was exacerbating these issues for them").

Meta also knew or should have known that the harms caused by its platforms would inevitably spill into schools, where those teens spend much of their lives and receive education, mental health, and other services. And, as discussed above, Meta had countless opportunities to warn the Districts through its partnerships with the National PTA, Scholastic, and its extensive outreach to schools. *See supra* § III.A.2.a. 4; *see also*, *e.g.*, Ex. 169 at 4962 ("A Parent's Guide to Instagram"); Ex. 170 at 2240 ("National PTA and Facebook Launch Digital Families Community Events"); Ex. 148 at 0829 ("Communications Plan: Back to School 2020…showcase how Facebook uses technology to … connect parents, students and teachers"). It said nothing. As one former Meta executive put it, "[t]his move was part of a broader pattern: when transparency threatens profits, Meta chooses secrecy." Ex. 8 at 122:22–123:12; *id*. at 105:12–107:19 ("The company doesn't provide any ability for us as a society to know"). By ignoring these warnings, Meta allowed its products to infiltrate classrooms, disrupt learning environments, and contribute directly to the youth mental health crisis now overwhelming schools nationwide.

(5) Meta's failure to exercise reasonable care

Meta argues that some of its users are, at most, addicted to and hurt by *the content* on Instagram and Facebook, not by the platforms themselves. That is a convenient defense, invoking Section 230. But it is not true. And there is substantial evidence, coming again from Meta's internal documents and research, proving that point. *See e.g.,* Ex. 466 at 3 ("our results cannot be attributed solely to the content of the post"). In this section, we turn to the dangers and harms inherent in Meta's platforms, Meta's failure to warn Districts about those risks, and its consistent prioritization of its profits over the remediation of known dangers to kids, including the Districts' students.

Before addressing those issues as to specific "platform features," we first take up Meta's failure to warn as to the risks "attendant to … platform construction in general." Dkt. 1267 at 14. Again, internal Meta documents demonstrate that its platforms *as a whole* are designed in unsafe ways that create harms including, but not limited to, addiction. Ex. 105 at 4015 ("Product features that are designed to exploit insecurity, or provide a dopamine rush (likes, notification, the pull-down-to-see, the infinite scroll, etc), to increase time spent, are inherently at odds with well-being and take away from people's ability to consciously focus[] on activities that add value to their lives."); Ex. 80 at 5384 ("Triggers" for "loss of control over time spent on Facebook" include "Notifications," "auto-play," "Fear of missing out (FOMO)," and the "Mystery of the algorithm" while "Life impacts" include "Loss of productivity," "Sleep disruption," and "Safety risks."). In 2018, Meta researchers even diagrammed the interaction among features that it hypothesized deliver "unpredictable, low value rewards and cues that promote frequent, automatic, undesired behaviors…in a way that is not aligned with users' intentions or preferences":



Ex. 211 at 8566, 8578 (defining "dark pattern" features as "interactions that are deceptive, or that trick you into doing something you didn't want to do" and "'addictive-like' design elements" as those "that promote repeated or poorly regulated behaviors that may have a negative impact on the individual over time."). Therefore, while we now proceed to a "feature by feature" analysis in keeping with the Court's analytical framework, this should not obscure the reality known to Meta (and undisclosed to the public): the design of Instagram and Facebook *writ large* present dangers to kids.

(a) Algorithm

Years ago, Facebook presented its users with a "News Feed" of posts from only their "friends," sorted in reverse chronological order. This is not how Meta's platforms have worked for well over a decade. Instead, Meta uses sophisticated machine-learning algorithms to determine what to show users and in what order. And that is no longer limited to posts from a user's friends—instead, it includes "unconnected" posts from other users across the platforms. Meta's change to a "ranked feed" led users "to spend 40% more time in the app each day," Ex. 212 at 8565, and has been an important revenue driver for the company. Ex. 219 at slide 5 (ranking "increas[es] the number of sessions and time spent when compared to chronological feed"); Ex. 217 at 9529 ("The ranking team has consistently driven +0.1% DAU per half for a year now!! It drives revenue! I've never seen anything like it!"). This success for Facebook led Meta to begin "building a ranked Feed" for Instagram in 2015. *Id.* The results were similar. Ex. 213 at slide 24 ("Introducing recommendations in IG Feed[.] Early tests show significant increases in consumption and time spent."). Today, both platforms use algorithmic ranking across a variety of "surfaces," including Facebook Feed, Facebook Notifications, Instagram Reels, Instagram Feed, Instagram Explore, Instagram Stories, and Instagram Notifications. Ex. 201 at 21.

Meta's ranking algorithms consider various kinds of user behaviors it calls "input signals." Ex. 201 at 15-16. These range from extremely narrow measurements ("How many users have clicked on the post in Instagram Explore to view it in full screen in the past hour," *Id.* at 126) to extremely broad ones ("How much time the user has spent in Instagram Explore in the past 84 days," *Id.* at 117). Input signals are fed into a complex AI model that predicts how a user will respond to a given piece of content. *Id.* 201 at 15-16. These predictions include, for example, "How likely the user is to reshare" the content, or "How likely the user is to watch more than 95%" of it. *Id.* at 117, 121. All of these predictions are weighted and summed (in what Meta calls its "value model"), and the post with the highest score gets displayed to the user. Ex. 985 ¶¶ 77-78. The algorithm then runs through this exercise for the next piece of content, and the next, and the next. Each video or photo is evaluated individually; their overall sequence is not considered. Ex. 14 at 309:21-310:18.

As this discussion demonstrates, Meta's ranking algorithms do not call balls and strikes on what content is "good" or "bad"—those determinations are left to separate machine learning models that

implement Meta's content moderation policies. Ex. 14 at 319:1-8 (Meta expert Birnbaum: content moderation and content recommendation "are separate sets of programs"); Ex. 220 at 155:3-13 (Meta expert Ferrara: "The two systems are distinct"); Ex. 986 ¶¶ 43-48 (Districts' expert Narayanan); Ex. 13 at 120:17-121:15 ("Facebook designs News Feed by considering how best to maximize its topline metric, not the particular content to show individuals."). Put another way, Meta's ranking algorithms "respond solely to how users act online." *Moody v. NetChoice, LLC*, 603 U.S. 707, 736 n.5 (2024).[4]

In any event, how Meta's ranking algorithms work is ultimately far less important than what they are designed to do. And what they are designed to do is maximize user engagement to increase revenue. That much is clear from Meta's documents. Ex. 215 at 8644 ("I feel like i'm reading totally different litigation and legislation than our org leaders… we literally are optimizing teens feeds for engagement."); Ex. 218 at 6686 ("the algorithm is absolutely tuned to maximize engagement in a maximally empirical, principle-less way. I am not sure that the youth legal people realize this and that it applies to teens just as much as gen pop. So the things contained in the lawsuits are somewhat hard to refute"); Ex. 216 at 6929 ("Teens are still developing, novelty-seeking and impulsive . . . Cool, so we probably shouldn't have ranking primarily driven by p(eyeballtime) and p(engagement), which is how it currently works."); Ex. 214 at 2552 ("core algo is…literally just mathematically optimizing for [impressions] and sessions for teens as with everyone else."), 2558 ("IG ranking is already different for teenagers…and FB ranking is different for YA [Young Adults]…in both cases it's to squeeze even more sessions out of them."). Even Meta's own algorithm expert has conceded this fundamental premise. Ex. 14 at 220:13-16 (Q: "So you're not disputing that Meta's ranking systems employ user engagement objectives, are you?" A: "No.").

Meta has not warned the Districts, or anyone else, that its engagement-optimized ranking algorithms can harm young users. Ex. 985 ¶¶ 245-253. But it knows this is the case, in several respects.

[4] The Supreme Court has declined to address whether such behavior is protected by the First Amendment. *Id.*; *see also NetChoice, LLC v. Bonta*, 2025 WL 2600007, at *13 (9th Cir. Sept. 9, 2025) ("some personalized recommendation algorithms may be expressive, while others are not, and that inquiry is fact intensive. *Moody* made that much clear."). The Districts expressly preserve for appeal their arguments that neither the First Amendment nor Section 230 protect the operations of Defendants' ranking algorithms (as distinct from their content moderation algorithms).

First, addiction is an inherent byproduct of algorithms designed to predict behavior to increase usage. Ex. 239A at 1761 (FB "optimizes for time spent … it just keeps people coming back even when it stops being good for them."); Ex. 13 at 108:21-24 ("advanced forms of AI that Facebook were using at that point to optimize for engagement" presented "risk that that might supercharge problems of addiction"), 67:12-17 ("given that the whole News Feed system is aimed at engagement, that's a set of repeated patterns of behavior – clicking, liking, sharing, and so on – the best way to increase engagement is addiction in some way"); *see also* Ex. 2A at 137:17-138:18 ("the design of the algorithm to promote sustained engagement … can play a role in drawing viewers into the experience for longer periods of time than they necessarily would want for themselves"); Ex. 230 at 0003 ("If algorithms favor content or functionality that encourages people to spend more time on Facebook, then it's possible that this will by its nature tap into addictive mechanisms" and "make addictive/compulsive usage more severe").

Second, addiction is a risk created by the algorithms' unpredictable delivery of rewards. Ex. 14 at 134:15-18 & Ex. 223 (Meta's algorithm expert acknowledging ten reasons for variability in user feeds); Ex. 108 at 5593-94 ("Maximizing for time-spent incentivizes bad ranking…we would be better off by adding some randomness or filler, so that the value is stretched out…I'd be surprised if it doesn't happen."). Meta understands that the variable delivery of rewards by its algorithms can lead to so-called "problematic use." Ex. 107 at 0201 ("triggers contributing to PU [problematic use] habits" include "Mystery of the algorithm - Uncertainty over if they will see posts from those they want; if they can find a post again later."); Ex. 225 at 3133 ("[w]hen rewards are unpredictable, dopamine release is increased"); Ex. 228 at 0075 ("When reward can't be reached reliably, people are incentivized to open the app more frequently – but their usage becomes habitual instead of meaningful."); Ex. 139 at 0127 ("Intermittent rewards are most effective (think slot machines) reinforcing behaviors that become especially hard to extinguish—even when they provide little reward"); Ex. 225 at 3135 ("[i]t may be a problem if Facebook seems rewarding based on the principle of unpredictability, while the inherent value of the reward is lacking."); Ex. 74 at 2152-53 ("people are binging on IG so much they can't feel reward anymore").

Meta knows (and has not warned) that these problems are particularly salient for its youngest users, as "[a]dolescence and early adulthood are characterized by unique sensitivities to rewarding stimuli, which manifest at the behavioral and neural level." Ex. 227 at 8021. Its "own product foundation research has

shown" that, because "a huge driver for teen behavior is the prospect of reward" and "[t]een brains are much more sensitive to dopamine," youth "have a much harder time stopping even though they want to." Ex. 224 at 0680; *id.* at 0676 ("IG has a pretty good hold on the serendipitous aspect of discovery…every time one of our teen users finds something unexpected their brains deliver them a dopamine hit. Teens are insatiable when it comes to 'feel good' dopamine effects."); Ex. 74 at 2152-53.

Third, because the very purpose of an engagement-optimized algorithm is to generate the most engagement—by whatever means necessary—it "very frequently creates/exacerbates integrity problems." Ex. 222 at 7835. It is a "a blind mathematical process with no principles." *Id*. The algorithm will promote whatever "gets lots of engagement," irrespective of whether that is what "users like, want, and value." Ex. 229 at slides 5, 7; Ex. 108 at 5591-93. This means, for example, that "a comment that you hate a thing can be seen as a positive signal leading content to get outsized distribution." Ex. 234 at 1252. What matters to the algorithm is the fact of the comment—it does not know, or care, what the comment says. *See also* Ex. 234 at 1258 ("even Mark himself has suggested that the anger reaction is a reasonable way to express that you don't like a piece of content, even as we count it as 4x more important as a less ambiguous 'like' for giving you more such content").

This creates additional harm for users. One former employee testified that Meta's ranking algorithm "was widely understood to increase the sharing of content that provoked feelings of shame or disgust or fear or pride—social emotions…that meant more user engagement on the system, and that generated more revenue for the company." Ex. 13 at 122:25-123:13. "[B]y defining value in terms of…clicks, likes, shares, plays – you ended up incentivizing repeated patterns of behavior on the tool that were in the end undesirable, and sometimes actively harmful for the users." *Id.* at 81:19-24; *see also* Ex. 233 at 7777 ("While FB Feed Relevance has optimized effectively for visitation and engagement, it has often increased the distribution of integrity problems–which can compound over time."); Ex. 232 at slide 10 ("Ranking models cause integrity regressions by increasing the prevalence of non-recommendable content.") & Ex. 3 at 455:15-25 (describing this as a "known consequence" of the "choice" to "prioritiz[e] recommending whatever [was] most engaging, most compelling"); Ex. 208 at slide 11 ("Meaningful Interactions are highly related to harmful content. Interactions such as clicks, likes, private shares, and comments are associated with sexually suggestive and hate speech content across IG.").

Meta knows this, and it knows the problem is the worst for teens. Ex. 214 at 2561 ("it seems to genuinely be the case that optimizing for sessions just naturally distributes more harm to teenagers."); *see also* Ex. 231 at slide 44 ("Bullying and Harassment (B&H), Suicide and Self Injury (SSI), and Violence and Incitement (V&I)…are all >2x higher prevalence for teens than the average user"); Ex. 236 at 7851 ("Feed Relevance consistently distributes SSI … to teenagers and nobody has noticed."); Ex. 245 at slide 11 ("SSI and ED [eating disorders] have a significantly disproportionate large teen audience."). Meta failed to fix this problem. *See* Ex. 2B at 759:14-760:12 (acknowledging that, 11 years after acquiring Instagram, Meta had not addressed this problem). And it warned no one. *See, e.g.*, Ex. 13 at 125:25-126:10; Ex. 235 at 3568 (2024: "imagine going out in public and telling people we rank their kid's facebook by what makes them consume the most content IDs lol" "this will be a public conversation and we are not ready for it" "and then when they hear what things keep their kid consuming the most content IDs … we're off to the races on a really fun conversation.").

Fourth, the engagement-optimized design of Meta's ranking algorithm is known to push users into "rabbit holes," feeds consisting of increasingly extreme types of similar content. Ex. 985 ¶¶ 212-218. Meta's would-be algorithm expert acknowledged that rabbit holing is "a plausible extrapolation from the operation of the system…the content in your feed becomes more and more focused on one specific topic, generally speaking." Ex. 14 at 363:21-364:8; *id.* at 365:20-22 (same). Importantly—and contrary to Meta's Section 230 defense—he acknowledged that this phenomenon arises from the operation of Meta's ranking systems, *not* from any piece or type of content. *Id.* at 366:11-16. Internally, Meta recognizes that its platforms can drive users, especially teens, into dangerous rabbit holes. Ex. 242 at 1808 ("someone feeling bad sees content that makes them feel bad, they engage with it, and then their IG is flooded with it"); Ex. 243 at 7939 ("if you start following borderline accounts or interacting with such content, our recommendations algorithms will start pushing you down a rabbit hole of more egregious content."); Ex. 241 at 3437 ("well-being challenges: … our ranking algorithms taking [users] into negative spirals & feedback loops that are hard to exit from."); Ex. 240 at 2551 ("people who are suffering from depression and self-harm go down IG rabbit holes, and explore functionality compounds this issue."); *see also* Ex. 247 at 0818 (Meta employee complaining, "I once got enticed by some snake video in watch, now all my recommendations occasionally have snake videos show up. It's terrible"); Ex. 256 at 24:15-25:8 ("once

the young people start to self-injure, they tell us that their feed…on Instagram suddenly gets full of this content."); Ex. 3 at 358:3-359:18. Meta also knows that its algorithm can drive adult predators into rabbit holes containing child pornography. *See* Ex. 2A at 282:15-23 & Ex. 325 at 0324 (testifying that investigators at Meta set up an Instagram test account that, within less than 24 hours, was filled with minor sexualization content and recommendations for child exploitation imagery).

Fifth, and relatedly, the structure of Meta's ranking algorithm can trap users in "filter bubbles" that validate and reinforce a single point of view. As with rabbit holes, Meta's algorithm expert acknowledged that filter bubbles are "a legitimate concern" that are "a kind of reasonable extrapolation from the way content recommendation systems work." Ex. 14 at 353:13-354:15. He described filter bubbles as situations where "you're not really able to see a bigger picture, that you're kind of trapped in a certain sense and you keep being reinforced for your beliefs and you don't see things that might contradict those beliefs." *Id.* Again, he acknowledged this phenomenon arises from the operation of Meta's ranking systems, *not* from any piece or type of content. *Id.* at 366:18-22. Meta recognizes that its platforms can reinforce, validate, and amplify negative beliefs, including negative beliefs about oneself. Ex. 246 at 7497 (internal comment liked 40 times: "we are far more likely to show people only content they will agree with (in the pursuit of metrics), creating an echo chamber. We are…biased towards metrics which ensures people's beliefs, whatever they may be, are validated and reinforced."); Ex. 248 at 5292 ("Preference Amplification is the *process* by which a user 'rabbit-holes' into a certain type of content. The process may or may not be driven by intentful user actions. PA can happen for any content"); Ex. 177 at slide 17 ("After connecting with accounts that include or create SSI-ED related content…recommendations for suggested accounts related to SSI-ED may be shown across surfaces.").

Rabbit holes and filter bubbles occur because Meta's algorithms recommend content on an item-by-item basis. Ex. 14 at 309:21-310:18. They do not understand when *sequences* of content may be problematic. Multiple witnesses acknowledged that Meta's algorithms can present harm to teenagers through the aggregation of photos and videos any one of which, on their own, might not violate Meta's content moderation policies. Ex. 2A at 221:9-12 ("The algorithm does sometimes push content that, in aggregate, could be seen as encouraging suicide"), 221:18-222:2 ("Instagram might promote content that, in aggregate, could be seen as promoting self-injury," including to kids); Ex. 249 at 7437 ("IG: lack

classifiers to detect and take action on problematic aggregated content on surfaces like Explore, hashtag pages, Reels."); Ex. 2A at 187:11-15 & Ex. 238 at 1685 ("when you zoom in, any one of those photos probably wouldn't be triggering, but in an Explore grid with all these photos stacked against one another, it's pretty overwhelming"); Ex. 2A at 191:1-20 (Explore grid for teenage girl consisting of dieting tips would not violate Instagram's policies but "definitely" could create a problem).

Once again, Meta's erstwhile algorithm expert made these very points. He acknowledged that Meta's ranking algorithms could lead "to a user seeing a lot of content…in a certain area," any of which "by themselves might not be a particular problem, but taken together in the aggregate" could be "potentially harmful." Ex. 14 at 330:14-333:16. Further, he admitted that this was a "potential extrapolative problem of content recommendation systems," a "serious problem," and one that "content moderation policies by themselves" were incapable of solving. *Id.*

As with the other problems with its ranking algorithm, Meta knew there was an issue and failed to warn anyone about it. Ex. 2A at 192:16-193:15 & Ex. 238 at 1686 ("[t]here are really meaty legal issues…largely by way of deceptiveness"). Nor did Meta fix the problem. Ex. 2A at 188:5-10 ("I don't think we had done enough work to, like, understand or mitigate the impact of aggregated surfaces."), *id.* at 188:21-24 ("due to just overall under-resourcing of this area, we certainly couldn't move as fast or as extensively as we would have liked, given the solutions that were available to us."), *id.* at 195:11-196:4.

All of this is particularly galling given that safer, reasonable alternatives to Meta's ranking algorithms *do* exist. *See* Ex. 986 ¶¶ 65-88 (Districts' expert Narayanan). For example, Meta could consider the expressed preferences of its users—by simply "ask[ing] users what they think is bad," Ex. 259 at 2710—rather than relying so heavily on behaviors that supposedly reveal users' unstated preferences. Ex. 985 ¶¶ 297-301; Ex. 247 at 0818 ("we are so short term focused … most of our AI systems are using behavioral user implcit [sic] signals vs Google which uses objective human ratings" / "survey based metrics don't have a level playing ground with behavioral metrics like sessions."). Meta could have given users the option to "fully reset their recommendations and ads," to escape from problematic rabbit holes. Ex. 177 at slide 47. Meta could utilize a technique called "listwise optimization" to mitigate rabbit holing and filter bubbles, Ex. 14 at 345:20-346:20, 352:22-353:2, and it understood that others in the "Industry [had] successfully" done so, including "successes in Baidu, Alibaba, Kuaishou, etc." Ex. 257 at 1306.

And, of course, Meta could default users to a pure chronological feed—which Meta has long been capable of doing, Ex. 281 at 7228, and which teens have long told Meta they'd prefer. Ex. 280 at 0439 (2020: "Repeatedly over the past year, across topics and workstreams, we've heard teens complain bitterly about Instagram's nonchronological feed…. Many say that if they could change one thing about IG, this is what they would change….Although the non- chronological feed fuels increased time spent on the site, teens say it makes it more difficult for them to find relevant info about their closer friends.").

Meta owes a duty of reasonable care to its youngest users and their caretakers, a point its own employees have recognized. *See* Ex. 2A at 187:2-5 ("the responsibility of making sure those are responsible recommendations falls to us."); Ex. 234 at 1247 ("We are Responsible for Viral Content"); Ex. 248 at 5291 (acknowledging there are "Users driven to harmful content by Instagram" who are "on the path to the worst experience, thanks to IG"). Yet it has failed to implement any of the safer available alternatives—for the simple reason that they would reduce engagement and thus revenue. *See* Ex. 2B at 383:14-385:11 ("a significant reason" the aggregation problem wasn't solved is that it was "generally pretty difficult to make safety changes that might impact growth or daily active users by any significant amount"); *see also* Ex. 10 at 73:12-74:25 (researcher who proposed change to reverse chronological feed "was laughed at the first time I brought it up" and told "we wouldn't be able to do something like that because it would lower engagement"); Ex. 13 at 56:14-57:7, 121:19-122:21 (researcher on Responsible AI team describing how "my team was folded and disbanded," which "demonstrated that the business model was prioritized over the safety of users"). Shockingly, technical documentation reveals that Meta allows changes to its ranking algorithm to regress safety *more* for teens than the population at large. *See* Ex. 392 at 7539 (changes to ranking algorithm allowed to regress safety by 0.24% for Teens but only 0.1% for the general population) & Ex. 14 at 372:1-5, 375:1-4, 381:6-384:12 (Meta's own expert conceding "it's intuitively plausible…that this would be less protective than it should be of teens. That's an intuitively plausible interpretation.").

Despite recognizing (and abdicating) its responsibilities to young people, Meta witnesses repeatedly testified that the company "did not warn parents or kids about the specific types of risks" posed by their platforms. Ex. 2A at 66:12-22. *See also* Ex. 168 at 144:6-15, Ex. 168B at 436:16-22 (testifying that Meta did not warn parents or their children use of their platforms could lead to addiction or

problematic use); Ex. 13 at 125:18-126:10 ("I was not aware of any efforts by Facebook to warn users of the potential impact of Facebook's ranking model design."); *See also* Ex. 997 ¶ 148 (discussing Meta's failure to be transparent with users and collecting testimony).

(b) Infinite Scroll

Meta's ranking algorithm works in tandem with another feature, infinite scroll, to "remov[e] stopping cues so that [a user's] brain doesn't wake up to catch up with impulses." Ex. 1070 at 41:16-18. The infinite scroll is exactly as its name suggests—a never-ending feed of content through which a user can thumb endlessly. This "transforms user engagement into a passive, habitual activity" that "keep[s] users in a loop of reactive consumption." Ex. 987 ¶¶116-17. The feature's inventor, Aza Raskin, has publicly stated that it "was one of the first products designed to not simply help a user, but to deliberately keep them online for as long as possible." Ex. 252 at 2. In deposition testimony in this litigation, he expressed that he "deeply regrets the unintended consequences" of the feature. Ex. 1070 at 69:4-10.

As early as 2017, Meta's internal documents show that the company's researchers were aware that the infinite scroll could promote problematic use of its platforms. *See* Ex. 211 at 8566 (including infinite scroll in a list of "addictive-like" features "designed to encourage a particular behavior that is not necessarily in the users best interests"); Ex. 105 at 4015 (describing features like infinite scroll as being "inherently at odds with well-being"); Ex. 10 at 71:14-21 ("infinite scroll allowed people to use the product more than they wanted or more than they had intended to and . . . could really exacerbate people's use of or problematic use of the products."); Ex. 225 at 3116 ("We should reduce cases where rewards are unpredictable or lacking in value, and reduce unintentional behavior, for example: …continuous content without breaks."). Despite concerns that "limitless content should come with cues for pauses, completion, or active choice to continue," Ex. 253 at 6108, and the ability to implement features that would address the harmful impact of infinite scroll, Ex. 254 at 5382-85, Meta's policy was to "give people control, but not in a way that hurts [engagement] metrics." Ex. 255 at 7644. Meta never warned its users that its infinite scroll feature could lead to problematic use of its platforms. Ex. 997 ¶¶ 162, 167.

(c) Screen Time Management

Meta knows that one way to solve problematic use is to simply limit how much time teens can spend on their platforms. *See* Ex. 12 at 131:1-136:4 (discussing several product features that Meta

considered developing to curb problematic use). But, during the relevant time period of this case, it has offered its users only weak, ineffectual, and sometimes outright inaccurate tools for this purpose. And it has done so expressly to avoid limiting growth, engagement, and (therefore) revenue.

As of March 2020, Meta's only "proof point" on "tech addiction/problematic use" was a feature telling users how much time they had spent on the platforms. Ex. 260 at 1780. ("there's no product work we've done in the last four years that comes close."); *id.* at 1784 (tool was "flagship launch[]" that was "intended to address problematic use"); Ex. 80 at 5380 ("Initial products were launched to give people more control over their use, though efforts were defunded end of 2019"). But it turned out this feature was tabulating users' time incorrectly. Ex. 260 at 1782 ("we're sharing bad metrics externally" which could result in "legal liability."). Meta considered "unship[ping] this without a replacement"—but it decided against doing so, concerned it otherwise "wouldn't have the credibility we now have in the social comparison/mental health partner space." *Id.* at 1780. In short, until 2021, Meta offered only one tool to limit problematic use, the tool was bad, and Meta left it up anyway for show.

In 2021, Meta considered launching a new control (called "Take a Break") to facilitate users seeking to limit their use of Facebook and Instagram. "Take a Break" nudges users to stop using the app after a 20-minute session. Meta knew this feature would be most effective if it was turned on by default because (1) "most users don't know about the tools" Meta has, and (2) "are not likely to go into settings to find them." Ex. 261 at 8248; *see also* Ex. 262 at 6152 ("[S]tats show us that most people just use whatever the default setting or filter is."); Ex. 3 at 199:8-12 (if safety features are "not searchable and they're not findable, it certainly limits their ability to help people."); Ex. 261 at 8256 (July 2019: "research shows that the tools we have are hard to find."); Ex. 987 ¶¶ 111-112 (collecting sources).

But the team responsible for "Take a Break" was unable to even *test* it as an "opt-out" default, because leadership viewed it as a "one-way door"— "features that once [launched], we basically cannot disable" because "it looks really bad." Ex. 12 at 104:6-105:23; Ex. 263 at 9807 ("Team was unable to test Take a Break opt-out default for new teens due to 'one-way door' principle."). Meta did not want to even crack this door open because of concerns that a default-on Take a Break would cause "regressions" (negative impacts) to "ecosystem metrics" (growth, engagement, and revenue). Ex. 264 at 7198; Ex. 265 at 1173 ("[W]e will: disregard [Take A Break] variants with material ecosystem regressions").

Meta launched "Take a Break" as an "opt-in" setting—meaning, users would have to know about the feature, then exercise the time and discipline to navigate the platforms' complex settings menu to turn it on. This was worse for user safety and worse for reducing problematic use. Ex. 12 at 164:15-24. But it was better for Meta's business metrics. *See* Ex. 266 at 45-46 (determining that opt-in version of Take A Break would have no "ecosystem impact" on "IG Time Spent," "IG Revenue," and "Stories Impressions," among other metrics). Meta knew that very few users would ever adopt Take a Break. Ex. 267 at 3662 (predicting that more than 99% of teen users were projected to ***not*** use Take a Break) *and* Ex. 113B at 441:21-442:2 (acknowledging same). And in fact this turned out to be the case. Ex. 982 & Ex. 270 at 3 (showing adoption rate for Take a Break of 0.45%). Nonetheless, Meta has publicly (and falsely) touted the effectiveness of this tool. *Compare e.g.*, *with* Ex. 268 at 3 (announcing the launch: "We're encouraged to see that teens are using Take a Break. Early test results show that once teens set the reminders more than 90 percent of them keep them on.").

(d) Barriers to Exit

Because Meta failed to implement effective tools to help users curb addictive and/or problematic use of their platforms, users often attempted to self-remediate by deleting their accounts entirely. *See* Ex. 107 at 0204 (describing how addicted users would take "extreme measures" by "deleting the [Facebook] app or deactivating [their account] to curb use."). But Meta intentionally designed its account deletion flow to be "fairly aggressive as we try to stop people from leaving," including telling users that their friends "will no longer be able to contact you through the site" and "requir[ing] users to choose a reason for why they're leaving." Ex. 104 at 0087; *see also* Ex. 1070 at 37:7-20 (describing Facebook "come-back emails that include the faces of all of your friends" which "harness[] the vulnerability [of] social reciprocity and social loss"); Ex. 987 ¶ 92 (depicting the thirteen screen deactivation flow for Instagram ending with a 30-day waiting period). This designs traps users in a proverbial "Hotel California"—where they can try to check out any time they like but few ever leave. *See id.* ¶93 ("This friction…create[es] barriers, wherein common actions that are potentially undesirable to the platform—such as deletion—are made unduly complex to suppress a user from achieving their initial goal."). Meta does not warn its users or the public that it designs its systems to discourage account deactivation and deletion.

(e) Notifications

Notifications from Meta's platforms are another known driver of social media addiction and problematic use for teens. Ex. 107 at 0201 ("notifications" among the "10+ triggers contributing to [problematic use] habits"); Ex. 105 at 4015 (notifications among the features designed to "provide a dopamine rush"); Ex. 300 at 3213 ("21% of teen WAU [weekly active users] say notifications make it harder for them to manage the amount of time they spent on the app."); Ex. 665 at 0086 ("notifications can sometimes make it difficult to do our best work, might dimmish our productivity, and/or could contribute to feeling of distress or being overwhelmed.").

Meta is aware of the well-settled science that, "when social media use displaces sleep in adolescents (via nighttime social media use), it is negatively correlated to indicators of mental health." Ex. 301 at 0423. "[P]reteens who get fewer than the recommended hours of sleep… were more likely to have brain networks that were less efficient, less able to adapt and rewire, and less resilient to stress." Ex. 302 at 1532; *see also* Ex. 1069 at 104:20-105:7 (agreeing that consistent sleep is beneficial to adolescent brain development); Ex. 981 ¶¶ 307-310 (collecting sources).

As such, it is jarring that Meta understood (though it did not disclose) that late-night notifications are a significant driver of problematic usage for teens. Its internal research reveals that "43.3% of [weekly active user] teens have at least 1 late night session"—meaning a session between 12am and 4am—and that "4.6% of [weekly active user] teens have 1 or more late sessions every night." Ex. 116 at slide 10. Further, "18% of [daily active users] start a session from a notification that was sent and clicked within 12 am and 4 am." *Id.* at slide 39. For teens who spend 25 minutes or more on Instagram every late night, over a third of that usage results from "a notification delivered within 5 minutes of starting a session on the app." Ex. 305 at 11. Clearly, Meta has a fair share of responsibility for the overall problem that "1/5 teens report waking up and checking social media in the middle of the night." Ex. 105 at 4007.

Since 2017, Meta's safety product teams have proposed launching products—such as "Off-Mode," (Ex. 105 at 4009), "Together Mode," "Face-to-Face mode," "DND [Do Not Disturb]," and "Quiet Mode" (Ex. 306 at 9) in response to feedback from young users who expressed that they "didn't want notifications to appear during certain periods of time, such as during school… [or] while sleeping." Ex. 307 at 6249; Ex. 306 at 5 ("[teens] would unanimously use a feature that limited notifications to certain times of the

day."). But Meta has consistently "shelved" these safety features, Ex. 306 at 9, or watered them down to the point of being useless. According to one former employee who helped developed "Quiet Mode," Meta was concerned that this feature would negatively impact metrics related to growth and usage, resulting in a version less protective of teen health. Ex. 12 at 142:19-143:2. Indeed, when this feature was launched in 2023 (six years after it was first proposed, Ex. 306 at 9), it was buried as an "opt-in" setting and used only by 0.26% of Instagram users. Ex. 308 at 3.

Meta has not wanted to limit notifications because notifications drive revenue. *See* Ex. 312 at 0649, 0666 (reporting that "50% of Android" and "⅓ of iPhone" daily active persons open Facebook "via clicking on a push notif" and that that "these sessions enjoy [a] disproportionately *higher rate of engagement*") (emphasis in original). Internal documents demonstrate that Meta "limit[s] the amount of information in notifications because we want people to come onto the site," Ex. 139 at 0117, and "optimize[s] when notifications are sent" to "improve engagement by sending notifications during the right time of the day." Ex. 1178 at 8981. Specifically, Meta's notifications research has determined that noon—or the middle of the school day—and 8-9 PM are "ideal times" to send push notifications to users' phones. Ex. 310 at 2378; *see also supra* Ex. 54 at 9227 (discussing "school blasts" that sent targeted push notifications to students within specific schools). Meta has even discussed the importance of using notifications to make their platforms "an urgent app for US (Girl) Teens" to "bring [them] back to the app[s]." Ex. 309 at 4381-82.

All of these decisions have been intentional and driven by "a very clear and strong message from Mark [Zuckerberg] that DAP (and specifically US DAP) is extremely important" and "a bigger concern [] right now than user experience." Ex. 311 at 7391-92. At no point has Meta informed its users, their caretakers, Districts, or the public at large that it intentionally engineers when notifications are sent, and what information they contain, to boost engagement on their platforms.

(f) Social Validation Metrics

Facebook was one of the first social media platforms to introduce the "like" button, a lightweight signal that allows users to express approval of someone else's post. Ex. 355 at 0962 ("2009 FB launches the like button"). Meta has known since at least 2015 that this feature can lead users to develop expectations of social validation and, by doing so, increase site visits and user engagement. *See* Ex. 356

at 5920 (Meta research indicating that "[s]elf-motivated site visits increased after contribution . . . which is consistent with the theory about feedback expectations."); Ex. 359 at slide 28 (stating that "Facebook's rewards are easily accessible" and that "expectation of quick responses and availability e.g., updates/counts of Likes" contributed to problematic use). Meta also knows that these dynamics are particularly salient for teen users. Ex. 357 at 1216 (likes and similar signals "encourage teens to continue engaging and coming back to the app" because "[a]pproval and acceptance are huge rewards" for teens). In 2016, Meta expanded the ability to "like" to numerous other emoji-based reactions—a heart, a laughing face, an angry face, and so on. Ex. 391

Meta knows that the visibility and quantifiability of social signals can impact the brain. Indeed, two former Meta executives went public acknowledging this very phenomenon in 2017. *See supra* § III.A.2.a.4; Ex. 137 (Palihapitaya: users "get rewarded in these short-term signals, hearts, likes, thumbs up" which leaves them "vacant and empty" and "forces [them] into this vicious cycle" because they "need it back"); Ex. 134 (Parker: "we need to sort of give you a little dopamine hit every once in a while because someone liked or commented on a photo or post… that's going to get you to contribute more content…it's a social validation feedback loop"). A robust body of literature (including Meta's own studies) supports these observations, showing that likes lead to increased dopamine activity. *See* Ex. 359 at 24 ("Instagram pictures with more vs. fewer 'Likes' activate the ventral striatum (VS).").

Troublingly, Meta's internal research also demonstrated that, by fueling users' "need for social validation," Ex. 98 at 25, and given the "visual nature of [Instagram]," Ex. 131 at 6265, visible and quantifiable reactions can lead users to develop negative feelings of social comparison. *See, e.g.*, Ex. 381 at 6913 ("like counts are a key driver of social comparison"); Ex. 382 at 0234 (we "can be pretty confident of a causal link between Like counts and social comparison."); *see also* Ex. 981 ¶¶ 130-131, 147, 149-157 (same); Ex. 4 at 54:20-55:10. Further, Meta understood that "social comparison is a key driver of well-being," Ex. 381 at 6913, and that it's linked to "multiple negative well-being outcomes (e.g., increased loneliness, worse body image, and negative mood or affect)." Ex. 383 at 7852; *see* Ex. 382 at 0235 ("Seeing more posts with high Like counts (1K+) was associated with feeling worse: more negative social comparison"); Ex. 127 at 42 ("at least some of this association is causal.").

This combination—the need for social validation and the quantifiability of popularity through likes—creates a perfect storm that leads many teens, and especially teen girls, to experience negative social comparison on Meta's platforms. Ex. 384 at 8545 (concluding that "negative social comparison… is a significant problem on Instagram" experienced by "68% of teen girls"); Ex. 131 at 6265 (problem is "historically high and growing—particularly among teens."); Ex. 1175 at 4450 (acknowledging that social comparison is "a big problem that IG core mechanics make worse"); Ex. 127 at 8 (reporting "**51%** of people experience social comparison on IG" and "33% of people have been feeling worse about themselves on IG for 'several months to a year'") (emphasis in original). One researcher described outside studies as "consistently point[ing] to appearance comparison as a key step towards internalization of the 'thin ideal' and other beauty ideals which often lead to body image issues and for some eating disorders/ suicide/ self-injury." Ex. 131 at 6265. Another described "negative appearance comparison and body dissatisfaction" as "early risk indicators that a teen is starting to struggle" with an eating disorder. Ex. 177 at slide 20; Ex. 119 at 1121 (relative to other platforms, Instagram "seen as having the highest impact" on "Body, Appearance Comparison"); Ex. 19 at 284:16-287:7 (teen girls told Meta that Instagram was "feeding the spiral" of negative appearance comparison, exacerbating low self-esteem and body image problems); Ex. 360 at 6269.

In response to these findings, Adam Mosseri announced in 2019 that Meta would be launching a product feature that "hid" like counts on users' posts, to help limit social comparison on Instagram. *See* Ex. 6 at 618:14-620:6 (discussing hiding likes as a way to "depressurize the experience" on Instagram); Ex. 385 ("Instagram will test hiding 'likes' in the US starting next week."). Meta researchers determined that this initiative, codenamed "Project Daisy," would make users "significantly less likely to feel worse about themselves." Ex. 1176 at 3256; Ex. 102 at 445:8-18; Ex. 357 at 1214 (Daisy "had a statistically significant impact on reducing the frequency of 'like' comparison for both teens and non-teens."). Even Mark Zuckerberg agreed that "there's something nice about not quantifying [likes]" because it "feels like people are just interacting rather than playing a game." Ex. 386 at 2022-23. Nonetheless, Zuckerberg remained "unconvinced rolling [Daisy] out completely…would be positive." *Id.*

After a series of tests in early 2020, Meta ultimately elected to walk back Daisy because, despite improving users' well-being, it was "pretty negative to FB metrics." Ex. 387 at 6684; Ex. 388 at 5702

(showing Daisy created an "ad revenue impact of minus 1% . . . and a [Time Spent] impact of negative 0.29%"); *see also* Ex. 251 at 1789 ("Instagram will need stronger evidence than it had for the February Mark review that Daisy is an important feature"). As a result, the feature was shelved until 2021, when it was added as an "opt-in" feature buried in settings. Ex. 389; Ex. 367A 164:19-20 (Daisy "was not a default"). Internal documents demonstrate that only 0.72% of users opted for these controls after they were launched (which was somehow "lower than [Meta's] estimates."). Ex. 388 at 5703. Between January 2023 and March 2025, only 3.48% of youth Instagram users turned on these settings. Ex. 390 at 3.

Meta's failure to hide likes by default is egregious given that it was widely regarded as "one of the clearest things (supported by research) that we can do to positively impact social comparison and well-being on IG." Ex. 380 at 1392; *see also id.* ("Daisy is such a rare case where a product intervention can improve well-being for almost everyone that uses our products."); Ex. 393B at 339:14-19 ("I do think Daisy is one of the cases where we have done a lot of research to show that there are benefits to launching it."); Ex. 394B at 431:2-15 (advocating for launch of Daisy because Meta had "concrete research" showing it improved social comparison and that "the positives of that outweighed the negatives"); Ex. 395 at 80:5-81:15 (agreeing that social comparison is the most important driver of well-being and that hiding like counts was the best way for Meta to reduce it).

Internally, members of Meta's safety teams lamented leadership's decision and complained that "[t]he bar for well-being (and social comparison) wins has been set too high to be achievable." Ex. 396 at 1460; Ex. 380 at 1392 ("[I]f we can't ship Daisy given the external comms around it and the supporting research, then I am doubtful that we can ship any broad product changes with the purpose of improving user well-being."). In its weakened form, the teams felt that Daisy did not meaningfully improve the well-being of Instagram's users. Ex. 2A at 168:9-12 ("I think that the version of Project Daisy that launched did not address the core issue of negative social comparison"); Ex. 380 at 1391 ("it will be difficult to make a much stronger case [for Project Daisy]"). By contrast, the growth team was unrepentant. *See* Ex. 397 at 7912 ("It's a social comparison app, fucking get used to it").

Meta has not informed the public of its decision to roll back the Daisy controls nor has it ever informed them that likes and other social validation metrics on its platforms contribute to problematic use,

negative social comparison, eating disorders, depression, and/or anxiety. *See* Ex. 997 ¶¶ 172, 175, 180, 182-83, 188, 191-92, 197, 201-02.

(g) Beauty Filters

As discussed above, Meta had extensive knowledge that many teens experience negative appearance comparison and associated mental health risks from use of Instagram. *See supra* at § III.A.2.a.4. Despite understanding these risks, Meta failed to address the visible quantification of "likes" on its platforms. But it did worse. It also began to add so-called "beautification" filters (aka "cosmetic surgery effects") to Instagram. Ex. 371 at 9481; Ex. 361 at 6561 (Kevin Systrom to Mark Zuckerberg, 2017: "We're building skin smoothing and face transforms in-house at IG"). These filters remove blemishes and change the shape of users' faces "in a way that must be achieved by plastic surgery." Ex. 371 at 9481. Meta knew that "teens are some of the biggest users of these filters (and teen girls in particular) on IG." Ex. 274 at 7020; Ex. 362 at 8117 (filters "are overwhelmingly used by teen girls.").

Meta also knew that allowing these filters meant "actively encouraging young girls into body dysmorphia and enabling self-view of an idealized face…that can result in serious issues." Ex. 362 at 8117; Ex. 129 at 1908 ("Sharing or viewing filtered selfies in Stories made people feel worse"); Ex. 983 ¶¶ 96-97, 101-107. A literature review conducted at Meta indicated that use of filters to manipulate selfies exacerbated the "risk and maintenance of several mental health concerns, including body dissatisfaction, eating disorders, and body dysmorphic disorder." Ex. 364 at 6297; Ex. 365 at 9941 ("the altering of selfies appears to be connected with negative impacts on…the person posting it…in terms of mental health, body dissatisfaction, and eating disorder behaviors."). Meta knew that "[t]hese extreme beauty effects can have severe impacts" on "the individuals using the effects" and that "[c]hildren are particularly vulnerable," as are "those with a history of mental health challenges [or] eating disorders." Ex. 366 at 135.

In October 2019, Meta temporarily banned cosmetic surgery filters while it consulted outside experts about the issue. Ex. 371 at 9481. A "significant majority" of these experts "confirmed [Meta's] hypothesis that [filters] had the potential to be very harmful, in particular to young people." Ex. 4 at 46:2-47:8; Ex. 366 at 135-36 (describing consultation "with 18 experts…A large majority, including individuals in the AR field, recommended prohibiting these filters, citing known impacts to body image and mental health"); *see also* Ex. 995 ¶ 156; Ex. 982 ¶¶ 269-293. These experts further reported to Meta that

beautification filters would be particularly harmful "amongst vulnerable populations (females, youth, those with a history of mental health concerns, etc.)." Ex. 371 at 9481.

However, Meta lifted its ban in May 2020. Mark Zuckerberg directed product teams to "allow people to create and share face-altering effects on Instagram" but "not allow these effects to be discovered or promoted in our Effects Gallery." Ex. 372 at 7843. One of the "Key Considerations" behind this decision was "competitiveness/growth"—specifically, a concern that banning such filters would have a "negative growth impact, simply because any restriction is likely to reduce engagement if people go elsewhere." Ex. 371 at 9481-82; Ex. 4 at 107:14-107:21; Ex. 374 at 7297; Ex. 2 at 163:15-165:1 (filters were very popular and "banning them would have felt like a competitive loss").

Meta's decision drew significant criticism, even from within the company. One senior executive emailed Mark Zuckerberg directly: "I respect your call on this and I'll support it, but want to just say for the record that I don't think it's the right call given the risks. As a parent of two teenage girls – one of whom has been hospitalized twice in part for body dysmorphia – I can tell you the pressure on them and their peers coming through social media is intense with respect to body image … I was hoping we could maintain a moderately protective stance here given the risk to minors." Ex. 374 at 7295; *see also* Ex. 5 at 89:20-90:5 & Ex. 375 at 3599 (different executive warning that reversal of the ban would mean "[w]e would – rightly – be accused of putting growth over responsibility."); Ex. 4 at 138:10-138:24 (by lifting the ban on cosmetic surgery filters, Meta "didn't prioritize or put first the well-being of teenagers and kids"); Ex. 368 at 4488 ("we have no idea what the long-term effects will be for this generation that has grown up comparing themselves to something that's … totally fake.").

Amazingly, Meta's poor decisions regarding Project Daisy and cosmetic surgery filters are not even the full extent of its failure to remediate the problem of negative social comparison on Instagram. Another research effort at Meta involved building an automated classifier to identify content presenting a high probability of negative appearance comparison.[5] If such content could be identified using AI, then

[5] Respecting the Court's order concerning Section 230, the Districts do not argue that Meta should be held liable for failing to detect and remove negative appearance comparison content. They offer this episode only to provide context regarding Meta's knowledge of the problem that its defective beauty filters made worse, and Meta's state of mind in failing to remedy that defect or warn the public about it.

Meta could avoid algorithmically recommending it to teens and others at high risk of negative social comparison. Ex. 12 at 40:18-41:16; Ex 1177 at 8073 ("we really do recommend A LOT of high-NAC content to teens!"); Ex 102 at 205:18-207:12 ("for 10% of teens, 57% of the recommendations in Explore are high-NAC"), *id.* at 218:13 ("the numbers for teen girls are higher."). Predictably—characteristically—Mosseri killed the project. Ex. 12 at 44:6-10 ("I know for a fact that he did not approve it."), 47:1-13 (the team was "pretty disappointed" because they "felt like they had a solution" to "a big problem" but "had their hands fully tied"), 68:2-13. The data scientist and data engineer who worked on the project quit: "they wanted to make a difference, and they felt like they just didn't have an opportunity to do that." *Id.* at 48:6-8.

In spite of this, Meta has never warned parents or their children that beauty filters contribute to problematic use, social comparison, eating disorders, depression, and/or anxiety. *See* Ex. 997 ¶¶ 165-67, 171, 174-75, 179, 182-83, 189, 191-92, 199, 201-02

(h) Account recommendations

Meta' platforms, in particular their account recommendation algorithms, have actively placed kids in harms' way, facilitating what Meta internally calls "IIC"—inappropriate interactions with children. Ex. 5A at 370:9-11. Meta has never warned the public about the breathtaking scale of its shortcomings.

Let's start with some statistics, revealed in Meta's documents. Instagram has a feature called "Accounts You May Follow," through which the platform recommends accounts to a user. Ex. 5A at 230:9-20. In 2023, this tool recommended to adult groomers "nearly 2 million minors in the last 3 months"—and "22% of those recommendations resulted in a follow request." Ex. 319 at 6182. The situation was just as bad in the reverse. According to an internal audit in 2022, Accounts You May Follow recommended 1.4 million potential IIC violators to teenage users, in a single day. Ex. 5A at 372:2-8 & Ex. 320 at 2028; *see also* Ex. 377 at 2578 (internal 2019 analysis estimating that there are 3.5 million profiles "conducting Inappropriate Interactions with Children" over Instagram DM).

The problem has been just as bad on Facebook. Facebook has an analogous account-recommendation feature called "People You May Know." Ex. 5A at 230:9-20. Meta knew that, at least through the end of 2023, "we were recommending minors to potentially suspicious adults and vice versa

in PYMK," Ex. 321 at 3522, and that this feature "was responsible for 80% of violating adult/minor connections," Ex. 329 at 3255; *see also* Ex. 995 ¶ 70 (expert report of Brooke Istook).

Short of actually fixing the problems with its account recommendation systems, Meta at least could have warned the public about the likelihood this feature would put kids in harm's way.

Further, Meta could have taken the simple and logical step of defaulting kids' accounts to the highest privacy settings, to prevent strangers from inappropriately interacting with kids. *See* Ex. 995 ¶¶ 150, 151, 155. Meta knew this was an appropriate solution. But, as detailed below, it dragged its feet for years before making the change—allowing literally *billions* of unwanted adult-minor interactions to happen in the meantime (including those facilitated by its account-recommendation algorithms).

Meta's culpability can be traced back at least to the second half of 2018. In August, Guy Rosen (Meta's Chief Information Security Officer) emailed Adam Mosseri (Instagram's CEO) naming the problem: "someone looks me up on Instagram/Facebook because I'm a cute girl, sends me a message (or even a friend request) and we start messaging. That is where all the bad stuff happens (From the "less bad" tier 2 sexual harassment like dudes sending dick pics to everyone; to the tier 1 cases where they end up doing horrible damage." Ex. 376 at 5891. Two months later, *The Atlantic* published an article titled "Instagram Has a Massive Harassment Problem," featuring an image of "the Instagram brand … in a Dumpster fire." Ex. 34 at 1 & Ex. 113A at 66:21-23. This article discussed the experiences of two 14-year-olds who described receiving expletive-laden comments about their appearance and death threats on Instagram. Ex. 34 at 2, 5. Said one: "My entire experience of high school was completely ruined by Instagram harassment. It's draining, it's anxiety producing." Ex. 34 at 2; Ex. 113A at 67:20-68:10.

At this point, any responsible company would have jumped into the breach to address a clearly significant youth safety issue. But it wasn't until close to a year later that Meta began to even study the problem, let alone implement a fix. In the fall of 2019, Meta designed a research plan acknowledging that its existing "privacy" setting on Instagram was, at best, a misnomer—and that "private" accounts could receive a heavy volume of unwanted interactions from strangers, through mentions, tags, and most importantly direct messages. Ex. 323 at 0532; Ex. 113 at 130:17-132:2, 227:13-19. Meta heard this first-hand from Instagram users, through a series of hour-long interviews conducted in Los Angeles. Ex. 323 at 0531 & Ex. 113A at 126:17-22. This qualitative research revealed that "participants were most sensitive

to unwanted contact and interaction" on Instagram, Ex. 332 at 6902 & Ex. 113A at 139:7-13, including "[u]nwanted messages that were sexual in nature," which "were particularly upsetting." Ex. 332 at 6903. One 14-year-old participant reported a particularly poignant story (which easily could have been included in *The Atlantic* article from a year earlier):

> Fake robot accounts DM you and say watch my free sex videos and the person sends it to 50 different kinds of people. This is kind of like, hey, get a $500 Visa gift card. I did that one time….I was real young. I thought I am going to get some money. It asked for my address. Plain 'ole little me, I was kind of ditzy. What it said to me after that was, we have your address now, we will stalk you. I was 11. I freaked out. That's probably one of the reasons I got a private account.

Ex. 332 at 6904 & Ex. 113A at 141:12-22.

In addition to articulating the problem (again) for Meta, participants were clear on the solution. "[A]lmost all participants thought that younger users and new users should be defaulted to private" settings. Ex. 332 at 6901 & Ex. 113A at 135:12-20. "Teens who began with a public account, and later switched to private, were very certain that their younger sibling and cousins should start with a private account." Ex. 332 at 6901. Based on this feedback, the Meta researcher in charge of the study made two recommendations: "default[] all teen accounts to private mode" and, "for private mode defaults, limit tagging, mentioning, and group DMs to connected accounts." Ex. 113A at 147:5-148:2, 149:11-25. "This is more in line with users' expectations and desires." Ex. 332 at 6907 & Ex. 113A at 149:1-9.

At this point, Meta could have implemented changes to ensure its safety tools lined up with user expectations—and, at least, warned users that Instagram's existing private mode was less protective than they expected. Meta did neither of these things. Ex. 113A at 145:25-146:25. Instead, it asked its growth team to study the "growth and engagement impacts of defaulting teen accounts into a private setting." Ex. 7A at 206:5-9. The growth team was pessimistic ("This will likely smash engagement, DAP, MAP etc") and skeptical ("What's the rationale for the push?"). Ex. 342 at 5175. So Meta put off the issue.

When it picked things back up, it again limited itself to exploratory studies. This time, in March 2020, researchers embarked on an investigation into what they called "Smart Defaults"—"a safer account model for teens" that would "[p]rotect teens from unwanted DMs" and "additional interaction[s]" in order "to keep them safe." Ex. 113A at 156:18-157:12; Ex. 333 at slides 1, 4. In line with the 2019 interviews in L.A., this research acknowledged that privacy settings for "mentions and tags are expected" and that

failing to make this change would "not match users expectations." Ex. 333 at slide 4; Ex. 113A at 159:24-160:8. Nonetheless, the research proposed not just the option of "default[ing] all new teens into private"—which it acknowledged would "keep more people safe from unwanted DMs, tags, and mentions." Ex. 333 at slide 16. It *also* outlined the option of "upsell[ing]" teens into an "opt-in" private experience—noting this would present "less risk to engagement metrics." Ex. 333 at slide 16. Indeed, the growth team had been called on once again. It had determined that a true private-by-default would result in a loss of 1.5 million monthly active teens a year. Ex. 7A at 377:8-14; *see also* Ex. 334 at 5 ("Your data pretty succinctly shows that taking away unwanted interactions via private default settings is likely to lead to a potentially untenable problem with engagement and growth.").

Faced with a clear choice, Meta chose profits over safety, informing its well-being researchers in April 2020 that "We will not create any new default interaction settings for private accounts." Ex. 334 at 1; Ex. 113A at 181:16-182:1. Worse, leadership asked these researchers to gin up a rationalization for this decision—"give us a story that frames our decision that we've already made" by "fram[ing]" this as a tradeoff between "safety vs value." Ex. 334 at 6. Privately, researchers were disgruntled. For them, this was "pretty much [the] worst-case scenario." Ex. 335 at 6621. They knew that making teen accounts private by default was necessary, because users "never actually went into the settings menu" to affirmatively opt-in to safety tools. Ex. 113A at 204:10-22; *see* Ex. 261 at 8256 (July 2019: "Current research shows that the tools we have are hard to find."); Ex. 2A at 72:14-16 ("most users…don't charge their default settings until they are in a moment of crisis."); Ex. 322 at 36 (acknowledging that only 0.08% of youth use Instagram's opt-in tool to limit comments and DMs from strangers). One researcher grumbled: "ISnt safety the whole point of this team? [sic]" to which another responded: "The point of this team is to prevent another Atlantic article about how shitty instagram is." Ex. 335 at 6621.

Despite privately expressing the belief that Meta should "man up and default people," *id.*, however, these researchers obliged their superiors. In a note published internally on April 3, 2020, they endorsed the idea "that we keep the current default interaction settings as-is," arguing that a change to the defaults that limited tags, mentions, and DMs from strangers would "prevent[] an important way that new users form positive connections with people they don't know" (and offering only a passing acknowledgment that such a change also "would significantly impact growth"). Ex. 1183 at 8558. A month later, one of the

two researchers who authored this note was promoted. Ex. 113A at 254:5-9.

Meta knew that placing teens into a default-private setting would have eliminated **5.4 million unwanted interactions a day** over Instagram direct message—including interactions like the harrowing ones described by *The Atlantic* two years earlier, and by interviewees in Los Angeles a year after that. Ex. 113 at 236:4-10; Ex. 336 at slide 17. It could have made a simple product change to keep kids safe. Having elected not to do the right thing, it at least could have warned the Districts, the public, parents, or kids of this danger on Instagram. It didn't do that either. Ex. 113A at 248:24-249:11, 253:19-254:3.

Five months after Meta's decision, the nightmare Meta could have avoided came to pass. Employees determined that inappropriate interactions between adults and minors had skyrocketed on Instagram Direct Message—to fully **38x** what the company was seeing on Facebook Messenger. Ex. 337 at 0623; Ex. 210 at 2309. Meanwhile, Meta was "automatically removing 90 times *less* amount of inappropriate interactions with children on Instagram than on Facebook." Ex. 2A at 289:19-22 (emphasis added); Ex. 337 at 0623. Meta understood their role in this problem—"[a]ctors take advantage of our tools on Instagram to find and inappropriately engage with children," Ex. 337 at 0623—yet they continued to say nothing to the public. Ex. 113A at 260:24-261:14. Moreover, Meta knew that the problem was being compounded by the launch of Instagram Reels in August 2020. By broadcasting "viral" videos of young teenagers to a wide audience of adult strangers, Reels put kids in harm's way. Ex. 3 at 55:25-56:18 ("young women who were producing content on Reels were experiencing unwanted interactions as a result of their content being shown to people that they did not wish to have it be shown to."), 57:25-58:9 (same). Researchers who suggested "filter[ing] out minor accounts from Reels altogether" to avoid this problem were ignored. Ex. 2A at 47:1-48:13 & Ex. 207 at 9534; *see also* Ex. 3 at 80:18-81:20, 106:17-108:6.

Internally, the controversy over Private by Default renewed. Once again, the growth team was asked to examine "the growth and engagement knock-on effects" of rolling out such a feature, including potential "friending losses" and "engagement declines." Ex. 7A at 212:2-12, 215:1-5; Ex. 343 at 8137. The growth team remained pessimistic ("we'll never really mitigate enough") and it remained skeptical ("This was never really about increasing safety" / "It was all for the pR wins" [sic]). Ex. 343 at 8137-38. In August 2020, it determined that private by default would result in a "2.2% hit to US DAU" and a "2.2% hit to global for teens"—"that is HUGE." Ex. 344 at 8209; Ex. 7A at 220:2-6; Ex. 345 (tabulating impact

on growth metrics); Ex. 346 at 1386 ("defaulting people to private on day 1 means that people just use the platform a lot less").

In August 2020, Meta's leadership was presented, once more, with a decision. Meta's policy, legal, communications, privacy, and well-being teams all recommended that private-by-default should "Launch Now," Ex. 347 at 6982, observing that this "will increase teen safety" and is "in-line with teen user expectations," "parent expectations," and "regulator expectations," Ex. 347 at 6983. The growth team—and only the growth team—recommended "Don't Launch (Now)," noting a likely impact of negative "2.2% topline teen DAP in 5 years." Ex. 347 at 6982; *see* Ex. 7A at 380:6-22; Ex. 7B at 703:9-19. Meta did not launch private by default in 2020. Ex. 7 at 236:2-3, 384:23-24; Ex. 352 at 3254 ("the growth impact was too high"); Ex. 353 at 5759 ("scrapped due to growth concerns"); Ex. 3 at 68:25-71:7.

In March 2021, the discussion was revived once again. Ex. 348 at 7502 ("Guess what, they're reviving private by default LOL" / "mygod"). This time—at last—leadership decided to launch a version of the feature. On March 16, 2021, Meta publicly announced that it would be "restricting DMs between teens and adults they don't follow." Ex. 338 at 2. Meta represented that, with this change, "when an adult tries to message a teen who doesn't follow them, they receive a notification that DM'ing them isn't an option." Ex. 338 at 3. Not surprisingly, however, there were some significant problems—beyond the fact that Meta's year-long delay had already facilitated 1.9 billion (=5.4 million x 365) unwanted interactions between adults and minors on Instagram.

"[C]ritically, private by default was originally supposed to be for all users;" what launched instead "was for new users under the age of 16"—not existing teen users, new 17- or 18-year old users, or new teen users who lied about their birthday. Ex. 2B at 748:5-11, 749:1-3, 749:21-750:4; Ex. 2A at 71:9-21; *accord* Ex. 7 at 241:13-14. Meta knew this was a problem. In July 2021 (after the launch), it learned from an internal survey that 13% of 13- to 15-years-old had received unwanted sexual advances on Instagram in the past seven days, overwhelmingly from strangers. Ex. 145 at 5033, 49; Ex. 168 at 307:10-25, 319:8-23 ("they are experiencing these issues…facilitated by the product design of Instagram").

Further, what launched wasn't actually a "true default setting." Ex. 2A at 74:13-75:2. As Instagram's former Head of Safety and Well-Being testified, "if you said to somebody, oh, this is an account that is private by default, I think an assumption would be, I sign up, I'm a teenager on the platform,

my account is private. I will go into settings and change that to be a public account if I want to." *Id.* at 73:8-13. Instead, what Meta launched was a "default preselected option on a screen," *id.* at 73:1-2—meaning, new users who self-identified as 16-or-under would "get a screen asking me if want to have a public account or a private account, and the private option is selected." *Id.* at 73:14-74:1. Meta knew this would result in fewer private accounts and knew this would be less protective of teen safety, but did it anyway because of growth concerns. *Id.* at 74:20-75:16. As one employee groused internally: "[this] is about looking good to regulators so that they don't block our under 13 year old IG version we are working on. Thats it. It has a terrible impact on teen engagement and retention and no detectable benefit on integrity metrics." Ex. 349 at 6694.

Moreover, while Meta's public announcement indicated that adults would not be able to DM minors with a private account, *see* Ex. 338 at 3, this belied the spotty execution of this tool. That was cast into stark relief a year later when a "sitewide emergency event" in February 2022 revealed that teens were *still* "receiving DM requests from unconnected adults, breaking [the] public commitment that we made." Ex. 339 at 2959; Ex. 113A at 271:17-23. Indeed, fully "50% of message requests" to teen users were still coming from adults. Ex. 339 at 2960; Ex. 35 at slide 15. The problem appears to have had numerous causes. First, "the logic to block DM requests from adults to minors" did not include users whose age Meta couldn't predict. *Id.*; Ex. 113A at 280:4-281:14. Meta knew this was dangerous because it knew groomers regularly claim to be minors—and "[o]ur age models are totally blind to this extraordinarily obvious exploit." Ex. 340 at 1116. Second, Instagram "still allow[ed]" teens to receive DM requests from "senders with no stated age." Ex. 35 at slide 14. Third, it allowed such DM requests from "senders outside the US who state they are teens, but may not be teens." *Id.*

Taken together, "unwanted DM requests" remained a serious risk issue for teens fully a year after Meta's "private by default" announcement, with Meta conceding (internally, never externally) that "there have been many gaps in fulfilling our promise." Ex. 35 at slide 15. Meta never corrected the representations in its March 2021 blog post to say that adults could still interact with minors on Instagram. Ex. 351; Ex. 113 at 290:25-291:11.

Meta did not apply private-by-default settings to all teen accounts (including existing users) until the end of 2024, well into this litigation. Ex. 2A at 71:14-20. Given the depressing history told above,

"Instagram Teen Accounts" warrants skepticism. Ex. 341 at 1. According to Meta's highly-publicized rollout, this feature places teen users by default into "the strictest messaging settings, so they can only be messaged by people they follow or are already connected to." Ex. 341 at 3. Of course, this is precisely the intervention that Meta's safety researchers had been advocating for (largely unsuccessfully) since the middle of 2019. A skeptical reader might wonder what has changed—other than five additional years of public pressure, scrutiny, and litigation. The discovery cut-off in this litigation has prevented the Districts from adequately investigating Meta's internal understanding of Teen Accounts' efficacy.

What the Districts can comfortably say is that Meta delayed defaulting teens into private accounts for years, prioritizing its own profits over preventing *billions* of inappropriate interactions between adults and kids. *See* Ex. 350; Ex. 354 at 8526 ("it was a huge growth hit like shockingly so" / "we did NOT do it for a very long period of time because of that" / "it's absolutely mad that it was so contentious" / "would [Mosseri] want any tom dick or harry being able to see all his kids' content, follow them etc? is he fucking nuts?"). Meta failed to warn parents, students, teachers, school administrators, or for that matter *anyone* about these issues. *See, e.g.*, Ex. 5A at 388:11-389:15; Ex. 6 at 239:5-21; Ex. 113A at 146:7-25, 260:17-261:14, 289:9-25, 291:3-292:9; Ex. 168A at 329:24-331:12; Ex. 2A at 315:9-17. It abdicated its responsibility.

(i) CSAM Reporting

Instagram and Facebook both allow users to report content that may be harmful. This is good in theory. But it only works if the reporting mechanism is effective, and if Meta does something with the reports it receives. Unfortunately, this is not the case for the most horrifying content that exists on Meta's platform, child pornography (sometimes called child sexual abuse material ("CSAM") or child exploitation imagery ("CEI")). Nor is it the case for repeat offenders who traffic children.

Meta went out of its way to stymie user reports of CSAM. As of March 2020, Meta did not have a specific way for people to report CSAM on Instagram (whether or not logged in as a user). Ex. 2A at 274:6-20; Ex. 244 at 5735 ("We don't have an in-app reporting option for CEI that I'm aware of."). When Vaishnavi Jayakumar, Instagram's then-new Head of Safety and Wellbeing, learned about this, "it was very surprising to me," "[b]ecause reporting CEI is sort of baked into most platforms." Ex. 2A at 275:23-276:2. Ms. Jayakumar raised this issue "multiple times," Ex. 2A at 276:7-8, but was told that it would be

too much work to build and, if Meta did so, it would be forced to review any reports that came in through that channel, meaning even more work. Ex. 2A at 275:4-276:15.

At deposition, Ms. Jayakumar testified that this approach did not strike her as putting the safety of kids first. Ex. 2A at 276:16-19. Contrary to what she was told, "[i]t is not a significant amount of product work" to allow users to report CSAM and CEI within Instagram—"you essentially add an additional option to the existing number of reporting…options that are out there." Ex. 2A at 277:8-12. Notably, Instagram at the time already allowed users to report a number of other far less serious violations directly within the ap. Ex. 2A at 276:21-277:2; *see* Ex. 182 at 8871 (photo of "IG User Reporting Flow" with dedicated options to report "spam," "intellectual property violation," and "promotion of firearms," among others, but no option to report CSAM). Without question, it would have been significant additional work for Instagram to *review* reports of CEI—but that is precisely the work Meta should have done to keep kids safe. Ex. 2A at 277:13-17. Meta not only didn't do this, it made its reporting flow intentionally complicated, to reduce the reports it had to look at. *See* Ex. 168A at 163:16-166:4; *see id.* at 161:10-22 ("only 1 percent" of people who have bad experiences on Instagram "finished submitting feedback through the reporting tool.").

Meta also failed to inform users that they could report CEI and CSAM sent via ephemeral message. "Disappearing Mode, which is also known as Vanish Mode on Instagram today," is a feature that "allows you to send messages to people and then have it disappear, you know, right after." Ex. 2B at 457:1-5. Because these messages are, by their very nature, ephemeral, users are unlikely to know they could be reported to Meta. Ms. Jayakumar testified that, "in order to protect minors who might be subject to inappropriate messages or grooming messages, we wanted to upsell and really promote the value of reporting within Disappearing Mode…to educate the user that even ephemeral messages can be reported and that they have a limited window of time in which to do so." Ex. 2B at 457:9-19. However, Meta expressly declined to undertake that educational effort—again, because "the team that would review the reports… wouldn't have the resources" to do so. Ex. 2B at 458:2-10.

Disturbingly, Meta's desire to look the other way extended even to the accounts of repeat predators. Ms. Jayakumar testified—and internal documentation confirms—that Meta had a "17x" strike policy for accounts engaged in the "trafficking of humans for sex." Ex. 2B at 459:9-23 & Ex. 199 at 7242

("This area is severely under-resourced given our SESTA-FOSTA liability."). Ms. Jayakumar explained: "That means that you could incur 16 violations for prostitution and sexual solicitation, and upon the 17th violation, your account would be suspended." *Id*. This, "by any measure across the industry, is a very, very high strike threshold." Ex. 2B at 460:9-13. Needless to say, Meta never told parents, the public, or the Districts that it doesn't delete accounts that have engaged over fifteen times in sex trafficking. *Id*. at 315:9-17.

(j) Age Verification

In September 2024 (well after these lawsuits were filed), Meta announced what it claimed to be a new suite of protections on Instagram, "Instagram Teen Accounts." Ex. 271 at 1. Meta has told the public that it will "automatically place teens in Teen Accounts, and teens under 16 will need a parent's permission to change any of these settings to be less strict." *Id*. at 1. It has assured the public that Teen Accounts provides "Built-In Protections for Teens [and] Peace of Mind for Parents." *Id*. at 2.

But the fundamental problem with Teen Accounts is the same problem that infects all of Meta's other loudly-touted youth safety tools, such as Private by Default. *Supra* at § 6(h). Meta does not know the ages of its users, and it cannot protect those it cannot identify. As whistleblower Jason Sattizahn explained in testimony to Congress, the "Teen Accounts announcement … elides the fact Meta does not have adequate data or processes to accurately track teen users"—though it "does have nearly a decade worth of research showing that Meta does not know the ages of its users." Ex. 273 ¶¶ 140-146.

Instagram did not even require new users to provide a birthdate during account setup until December 2019. Ex. 274 at 7020. It wasn't made mandatory for existing users until 2021. Ex. 195 at 7; Ex. 113 at 79:23-80:2. As a result, for years, Meta did not even have self-represented age for an enormous fraction of Instagram's user base. Ex. 240 at 2552 (Apr. 2020: "we also only have age for 55ish percent of people"). In any case, Meta knew that simply asking users to input a birthday would be ineffectual. Ex. 9A at 72:8-73:12 ("pretty straightforward for people to enter a different age"). And it was. Meta is aware that at least "30%-40% of teens are registered as being >18." Ex. 25 at 13. Some internal estimates put this number even higher—at 60% of teens. Ex. 194A at 107:12-108:2. Further, data produced in this case reveals enormous "blind spots" in Meta's understanding of the ages of its users. *See, e.g.*, Ex. 282 (users with "unknown" predicted age: 142 million (Feb 2018), 162 million (Feb 2019), 182 million (Feb 2020),

204 million (Feb 2021), 216 million (Feb 2022); *id.* (in Feb 2018, it was "unknown" whether 16.2 million monthly active users were "teen" or "non-teen"); Ex. 283 (authenticating data). Needless to say, "Built-In Protections for Teens," Ex. 271 at 2, won't work if Meta can't even identify who those users are. *See* Ex. 2B at 476:6-17, 730:23-732:16.

Meta's lack of understanding of the ages of its users is problematic for an additional reason—it leaves Meta unable to comply with its obligations under federal law. Meta is required to observe numerous data privacy safeguards for any user under the age of 13 years old. *See* 15 U.S.C. § 6501 *et seq.* ("COPPA"). In this litigation, Meta has admitted it does not observe any of these protections. Specifically, "since 2012, Meta has not obtained verifiable parental consent for individuals under the age of 13 to use Instagram," Ex. 184 at 9, it "has not provided parents means to review personal information for individuals under the age of 13," Ex. 184 at 11, and it "has not provided parents a means" to "refuse to permit the further use of personal information collected from their children under 13." Ex. 203 at 11-12. *Compare id. with* 16 CFR §§ 312.3-5 & 16 CFR § 312.6.

Meta's justification for its noncompliance is that it "does not permit individuals under 13 on its platforms." Ex. 190 at 4; *see* Ex. 191 at 1 (same). This would make sense if it were true. But it is not. In fact, for nearly a decade, Meta looked the other way when it came to underage users of its platforms. Multiple company witnesses acknowledged that Meta is "generally aware … that there are people who are under the age of 13 who lie about their age in order to get around the controls that different apps have." Ex. 11A at 425:16-426:17 (Zuckerberg); *accord* Ex. 286A at 110:20-111:3 (Antigone Davis); Ex. 113A at 216:20-217:24 ([redacted]); Ex. 287 at 3666 ("definitely kids this age on IG"). But Meta's knowledge is not just generic. As early as 2017, Meta knew that "FB and IG currently predominantly function as consumption platforms for tweens." Ex. 317 at 4804. It knew that "[t]eens typically join Instagram when they are in middle school at age 11-12." Ex. 275 at 6272-6273 ("timeline of Instagram usage"). It even observed that "Instagram Sells Itself: In 6th grade, age 11-12, teens get first phone and are 'invited.'" Ex. 314 at slide 26; *see also* Ex. 296 at 4832 (quoting "11yo girl": "More people my age are into Instagram."). By 2018, Meta knew "there were 4M people under 13 in 2015 on IG" which "represents around 30% of all 10-12 years [sic] old in the US." Ex. 192 at 3546. It knew that "45[%] of children 9-12 say they use Facebook daily" while "40 percent use Instagram." Ex. 193 at 5502.

These numbers are hardly an accident. They are the foreseeable result of Meta's persistent interest in capturing users as early as possible—because "Tweens and young teens are Facebook's best opportunity to win future generations of people." Ex. 315 at 1198; *see also* Ex. 313 at 5642 ("Tweens (approximately age 10 to 12) are special. People who join[] Facebook as tweens have the highest long-term retention out of all age groups."); Ex. 316 at slide 4 ("We have definitively established tweens as the highest retention age group in the United States.").

That interest has manifested in extensive research into the psychology and digital behavior of "tweens." *See, e.g.*, Ex. 297 at 8628 ("Objectives: To unearth digital behavior of 8-12s"); Ex. 294 at 1309 ("Tweens are herd animals, and want to find communities where they can fit in."); Ex. 296 at 4809 ("Tweens are already on social media."), 4821 ("The majority of 10-12 year olds have at least one social media account."), 4845 (outlining research plan for social media users "approx age 10-13"). It has also included exploratory missions into new products designed for "users as young as 5-10." Ex. 293 at 1556; *see* Ex. 295 at 4410 ("Exploring playdates as a growth lever for Messenger Kids"); Ex. 298 at slide 13 ("**5-10 year olds are our best initial U13 segment,** allowing us to build[]on FB parent strength to grow a portable graph for young kids who will become future teens.") (emphasis in original).

Internally, employees were nothing short of disgusted at Meta's express interest in the tween market: "oh good, we're going after <13 year olds now? [Z]uck has been talking about that for awhile… targetting [sic] 11 year olds feels like tobacco companies a couple decades ago (and today). Like we're seriously saying 'we have to hook them young' here." Ex. 299 at 4290.

Meta has not acknowledged the widespread use of its platforms by tweens. Instead, it has denied the problem, "avoid[ing] conversations that highlight <13 yo with IG accounts," Ex. 279 at 8207, and declining to "ask [users] if they're under 13 because it means we'd have to delete their account," Ex. 285 at 3406; *see also* Ex. 168A at 183:25-184:4 (testifying Meta had a "don't ask, don't tell" policy when it came to underage use of Instagram from 2012-2015). When pop star Jojo Siwa told Adam Mosseri she had started an Instagram account at the age of 8, Mosseri's response was that he "did not want to know." Ex. 292 at 15. (Mosseri subsequently copped to the Senate that this was a "missed opportunity." *Id.*). Meta has even ventured beyond denial into outright deception, publishing "community standards pages mentioning we age gate"—though, in reality, this was "not implemented in IG." Ex. 276 at 5781; *see id.*

("we don't even have age on IG for a big chunk (last figure I saw as ~30%) of IG users...I wonder if the authors of those web pages know that?").

Meanwhile, Meta failed to provide users with good tools to report under-13-year-old users. *See* Ex. 196 at 6516-17 ("The [under 13] reporting flow was pretty bad….It was obviously structured to deter any reports."); Ex. 2A at 98:25-101:13 (reporting flow "very clunky" and resulted in "a lot fewer reports than otherwise"); Ex. 197 at 8850 ("currently there is no way for users to report U-13 users on Facebook in app"). And it failed to review millions of the reports it did receive—between 2020 and 2021, it maintained a backlog of between 450,000 to 2.7 million suspected under-13-year-old accounts. Ex. 199 at 7246; Ex. 200 at 6780; Ex. 288 at 2800 ("big backlog and demand is outpacing supply"). Removal of these accounts was deemed "not a priority," despite the fact that it meant "our basic COPPA compliance is at risk." Ex. 198 at 6543; Ex. 2A at 90:3–4 ("there weren't staff allocated to reviewing and addressing the backlog"); *see* Ex. 174 at 5153 ("The fact that we have age limits which are unenforced (unenforceable?)…makes it difficult to claim we are doing all we can.").

Meta could easily implement processes to find and remove under-13-year-old users (to comply with federal law) and correctly identify under-18-year-old users (to give its anemic safety tools at least some shot at success). Numerous options were, and remain, available. *See* Ex. 171 at 6015 ("if we used age classifiers we could detect under 13s and kick them off the platform."); Ex. 272 ¶ 14 (declaration of Sattizahn); Ex. 989 ¶¶ 217-264 (Estes opening report). Indeed, Meta has used such tools for the benefit of its advertisers. Ex. 276 at 5780 (2021: "in the past, the only requirement for age-gating on IG was by advertisers, around brand safety. That's why the relevant infra and code are owned by the IG Ads team"); Ex. 2 at 464:10-21 ("we sometimes…[would] try and predict the age of that account, to decide what kind of ads might be most relevant to them."). Meta has also used ID verification in the European Union, to prevent *adults* from misrepresenting themselves as *teens* "to get an ad-free experience." (The EU prohibits ad targeting to minors.) Ex. 290 at 1446. While Meta could easily use ID verification to "identify[] if a user is a teen," it did not do so because of "the potential revenue impact it can cause." Ex. 291 at 1601. As one former employee observed, if Meta can conduct age assurance "for business purposes," it "should also be able to do that for integrity purposes." Ex. 2B at 464:10-21.

But Meta's executives have stifled any efforts to implement new methods of age assurance to remove underage users or protect kids. *See e.g.*, Ex. 174 at 5145 (Bickert, 2019: "Would we really want to have public statements from our expert that Instagram needs to have age gating[?]"); Ex. 198 at 6543 (2020: "Adam Mosseri will not support upselling age collection for existing users because he sees it as intrusive data collection."); Ex. 284 at 6683 (2020: "no explicit decision on improving our existing age models. The leadership decision was limited to NOT building an u13 classifier") Ex. 199 at 7245 (2020: "Overall we are quite vulnerable when it comes to age management on our apps. To some degree that is because leadership has decided [to] live with certain risks at least in the near term."); Ex. 272 ¶ 6 (describing "sudden cancellation" of age assurance research project in 2022). Meta simply will not do so unless it is required. Indeed, after Arkansas passed a mandatory age verification statute, Meta built an age verification tool that was "ready for launch"—but it "paused" its implementation when the law changed. Ex. 289 at 1568-1569.

(k) Parental Controls

Meta's delays in launching parental supervision features on Facebook and Instagram have been unacceptable. In 2009, Facebook's founding engineer sent Mark Zuckerberg and Chris Cox an email proposing parental supervision tools. "I would like to see us add an opt-in feature, which would allow a Facebook user (child) to designate another user (the parent) to have certain auditing rights and limiting controls over the child account," he wrote. Ex. 318 at 5130. "The dynamic that this creates is to give parents an opportunity to act as parents on Facebook as they would in other dimensions of their children's lives … which I believe is the only scalable and effective way to address the issues of minors on Facebook." In reply, Chris Cox (now Meta's Chief Product Officer) responded that this "doesn't immediately strike me as high-leverage as much as nice to have" and "seems like a bunch of work for things that the parent could do manually." *Id.*

Meta did not make parental controls available to Facebook users until 2023—fully 14 years later. *See* Ex. 184 at 7 ("Meta admits that parental controls, including tools developed for parental supervision of Teens, were first made available to users on Facebook in 2023."). It has offered no explanation or justification for this delay. Nor has it explained why only 0.15% of Youth users are enrolled in these parental supervision tools—meaning, 99.85% are not. *See* Ex. 168B at 581:17-582:10 (Béjar: making

parental supervision opt-in meant "that the feature would not be adopted and then as such would not be effective as a safety feature"). The timing and adoption rates speak for themselves. Facebook's parental supervision tools have been nonexistent or ineffective during the relevant time period of this litigation.

The situation is no better with respect to Instagram. Meta only launched parental controls on Instagram in 2022, ten years after it acquired the platform. Ex. 184 at 6 ("Meta admits that parental controls, including tools developed for parental supervision of Teens, were first made available to users on Instagram in 2022."); Ex. 2B at 736:11-739:20. Further, "[t]he initial version of parental supervision required both parties to agree to supervision, so the teen or the parent could disable supervision." *Id.* at 737:21-24. Not surprisingly, adoption rates were as dismal as they were on Facebook—with only 0.38% of youth enrolling in supervision on Instagram from November 2023 through March 2025. Ex. 322 at 39; Ex. 308 at 3. Meta ultimately took an "about-turn from th[is] position" with the launch of Teen Accounts, *id.* at 738:16-21—and, without a hint of irony, even sent Instagram's CEO on the *Today* show to tout the importance of safe defaults. Ex. 328 (Adam Mosseri, Oct. 14, 2025: "I've got three kids, I know parents are busy, which is why it's so important that the defaults are safe."). But the fact remains that it took 12 years after acquiring Instagram for Meta to adopt parental supervision tools that a teen couldn't just disable on their own. *Id.* at 739:15-25. And even today, the parental supervision tools on Instagram Teen Accounts are only as good as the age verification behind them. Those tools remain broken. *See supra* § (6)(j).

At this point in the brief, the reader will hardly be surprised to learn *why* Meta delayed for as long as it did in releasing parental supervision tools. Parental involvement in teen accounts was seen as a threat to growth—limiting teens' content production, sharing, and network development. Ex. 327 at 1975 ("We should be thinking about how parents being on Instagram might effect [sic] graph management and teen engagement over time."); Ex. 298 at slide 14 ("Adults are Toxic: Unless we allow teens to share separately from adults, the rest of the world will inevitably follow the decline of teen engagement in the US and EU."). More succinctly: "Parents are a killjoy." Ex. 324 at 8318.

Needless to say, this is no excuse for depriving parents of the tools Meta knew they wanted, needed, and were entitled to in order to keep their kids safe. *See, e.g.*, Ex. 98 at 9218 ("parents can't understand and don't know how to help" because "social media has fundamentally changed the landscape of adolescence"); Ex. 177 at slide 11 (parents "do not have the ability to take direct action to protect their

teens from potential harm; >65% of parents say they want an increased sense of control" on IG); Ex. 326 at slide 3 ("Caregivers want support and education from Instagram on features and tools that can help them keep their teens safe.") (emphasis in original).

(6) Meta's misrepresentations and manipulations

Meta's failure to exercise reasonable care is compounded by a series of outright lies—which prevented even the most vigilant administrators, teachers, parents, and students from understanding and heading off the dangers inherent to Instagram and Facebook. We have already discussed Meta's false statements to Congress about its knowledge of addiction and anxiety, as well as its publication of false data underreporting the prevalence of problematic use. *See supra* § III.A.2.a.4.

Meta also lied to Congress about its goals—claiming, falsely, that it has not set out to increase the time users spend on its platforms. Ex. 11A at 171:12-16 (Q to Mr. Zuckerberg: "So three times before Congress you've testified under oath that you don't give your teams goals of increasing the amount of time people spend on your platforms, correct?" A: "Yes.") & Ex. 149, Ex. 150, Ex. 151 (video of testimony). In truth, Meta has not only set quantitative "time spent" goals, it has set those goals *for teen users in particular*. Ex. 152 at 2613-14 ("Q1 goals" in 2014 include "Time spent. Grow by 5% for H1 (across all interfaces, including Instagram"; Zuckerberg: "The most concerning of these to me is time spent."); Ex. 153 at 3741 & Ex. 11A at 178:15-17 (re: 2016 document - Q: "the second goal is time spent increases by 12 percent, correct?" A: "Yes."); Ex. 154 at 4688, 4692 ("2017 Teens Strategic Focus" document sent to Zuckerberg: "Strategic goal: Time Spent Share") & Ex. 11A at 181:17-182:12 (Q: "And there's an example of that underneath, right, 15 billion minutes per day across the Facebook family?" A: "Yes."); Ex. 155 at 2304 (H2 2017 Instagram "Goals" include "US Teen DAP Accounts" of "12.1M" and "Time Spent" of "28 mins") & Ex. 11 at 184:15-186:9; Ex. 156 at 0881 (H1 2019 Facebook document: "Time spent (per DAP) is showing sustained growth and is ahead of H1 goal"); Ex. 157 at 3524 (outlining targets for Instagram of "40 minutes per DAU * 1.5B DAU" in 2023, growing to "46 minutes per DAU * 1.95B DAU" in 2026); Ex. 158 at 0347 (Adam Mosseri: "My stretch goal was to get on track to pass TikTok in terms of time spent worldwide"); Ex. 159 **at** 7078 & Birnbaum Dep. 269:22-270:20 (Meta's expert: Q: "as of the second half of 2023, Facebook was measuring business value through a metric called revenue weighted time spent, yes?" A: "That seems to be the case."). As one Meta employee put it: "No

one wakes up thinking they want to maximize the number of times they open Instagram that day. But that's exactly what our product teams are trying to do." Ex. 226 at 1686; *see also* Ex. 330 at 1441 (LinkedIN founder Reid Hoffman to Mark Zuckerberg: "I've chatted with a few ex-FBers…they all (to a person) told me that FB culture is emphasizing sessions (#, length) over everything else and that they felt that this interpretation of 'emphasizing profit over people' was accurate").

Meta has also lied repeatedly about the existence of dangerous content on its platforms. That includes images and videos glorifying eating disorders, self-harm and suicide, and it even includes child pornography (which, the Court needs no reminding, enjoys no First Amendment protection as it "is intrinsically related to the sexual abuse of children," *New York v. Ferber*, 458 U.S. 747, 759 (1982)). To be clear, the Districts do not contest that Section 230 provides Meta with the latitude to publish, or de-publish, whatever content it pleases on Instagram and Facebook (although the possession and distribution of child pornography is a federal crime, *e.g.*, 18 U.S.C. § 2252). But nothing in the law gives Meta the right to lie about its choices to children and their protectors. *See, e.g.*, *Barnes v. Yahoo!, Inc.*, 570 F.3d 1096, 1107 (9th Cir. 2009), *as amended* (Sept. 28, 2009) (breaking a promise is not protected by Section 230); *Doe v. Internet Brands, Inc.*, 824 F.3d 846, 851 (9th Cir. 2016) (failure to warn is not protected by Section 230).

In early 2019, Meta responded to the death of Molly Russell, a fourteen year old British schoolgirl who, according to an official investigation of the London coroner, "died from an act of self-harm whilst suffering from depression and the negative effects of on-line content." Ex. 161 at 2; *see id.* (coroner: "The platform operated in such a way using algorithms as to result, in some circumstances, of binge periods and images, video clips and text some of which were selected and provided without Molly requesting them."). Meta issued a statement stating, in part, "We do not allow content that promotes or encourages eating disorders, self-harm or suicide and use technology to find and remove it." Ex. 160 at 23; *see* Ex. 180 (Feb. 2019 statement by Mosseri). Meta also updated its public-facing policies to state, "we do not allow people to intentionally or unintentionally celebrate or promote suicide, self-injury, or eating disorders" and "[w]e remove any content that encourages suicide, self-injury or eating disorders…and any self-injury content which is graphic regardless of context." Ex. 162 at 1 & Ex. 163 at 1; Ex. 2A at 218:16-20. Meta executives also appeared before the Senate, testifying, "[w]e don't recommend…eating disorder-

related content to people of any age," Ex. 164 at 27 (Adam Mosseri, Dec. 8, 2021) and "we do not direct people toward content that promotes eating disorders…We actually use AI to find content like that and to remove it." Ex. 165 at 29 (Antigone Davis, Sept. 30, 2021). These are lies.

Start with eating disorder content. It is true that Meta relies, in part, on AI tools (which it calls "classifiers") to detect content that violates its eating disorder policy. Ex. 166 at 16-17. But it is false that Meta uses AI to remove that content. In this litigation, Meta has admitted "under penalty of perjury," *id.* at 35, that its eating disorder classifiers "do not currently perform automatic deletions," *id.* at 16. Further, "therefore there is no threshold value for autodeletion," *id.*—meaning, even if Meta's AI tools determine with 100% confidence that a photo or video violates its eating disorder policy, Meta will *not* automatically delete it. *Id.* at 16-17; Ex. 985 ¶ 106 (expert report of Prof. Arvind Narayanan). Not surprisingly, and contrary to Meta's representations, Instagram and Facebook are awash in content that promotes eating disorders. According to a 2019 survey of 30,000 Instagram users (which Meta never disclosed), 18% reported viewing content that promotes eating disorders or unhealthy weight loss in the previous week. Ex. 168A at 123:25-124:5; Ex. 101 at slide 6. As of 2020, it remained the case that "it's easy to find violating ED accounts in a couple of taps. When I search for 'thinspo' or 'proana' the accounts shown are overwhelmingly violating." Ex. 179 at 6755. Meta knew this was "us not delivering on the promise made by Adam in February 2019." *Id*. at 6751.

The situation is just as bleak with respect to content promoting self-injury and suicide. Again, Meta uses AI classifiers to detect this content. But again, it does not use those classifiers to take reasonable and prompt action. Meta's under-oath discovery responses reveal that, even if its classifiers are 93.99% certain that photos and videos glorify self-harm, Meta will not automatically remove that content from its platforms. Ex. 166 at 13 ("Threshold value for autodeletion: 0.94."). Internally, Meta employees understand that "violating content" can "score[] significantly lower than th[is] established threshold." Ex. 182 at 8895 (admitting this is a "root cause" of the "problem"). Again, the results are predictable: "There's going to be SSI [suicide and self-injury] content on IG and FB for the near future. The existing classifiers are not hard to game." Ex. 171 at 6015 (internal chat from 2019); Ex. 172 at 0021 ("Suicide Video Producers and Suicide Promoters have some very high reach accounts"). As of 2018, Meta knew that, "[o]n a daily basis, 5.1 million users, or 0.75% of DAP, are exposed to SSI-related content." Ex. 178 at

4409. As of 2021, it knew (through a survey of nearly 250,000 users) that 6.7% of respondents were exposed to self-harm or suicide on Meta's platform in a given week. Ex. 145 at 5049 & Ex. 6 at 506:6-507:11; *see also* Ex. 146B at 441:7-11 (survey was "massive even for Meta survey standards"); Ex. 3 at 241:15-18.

This survey also found that the users most exposed to this content are the youngest—with 8.4% of 13- to 15-year olds reporting having seen "someone harm themselves, or threaten to do so, on Instagram" in the past week. Ex. 145 at 5033, 5049. It was "well known among the well-being team at Meta … that suicide and self-injury and self-harm content was easily accessed on Instagram by kids." Ex. 168 at 346:13-347:1. "We know that SSI and ED have a significantly disproportionate large teen audience." Ex. 177 at slide 11 (Oct. 2023). "Everyone involved with SSI and eating disorders," including "producers" and "reporters," "are more likely to be in early high school or late high school." Ex. 172 at 0030.

Meta executives knew there was a problem. Ex. 174 at 5152 (Monika Bickert, 2019: "IG's product is essentially violating our content policies by displaying self-injury images without captions."); Ex. 175 at 5341 (Nick Clegg, 2019: "I said publicly…that we would do 'whatever it takes' to address the legitimate concerns raised….This lack of connection between what we commit to publicly and what we do internally really worried, and slightly shocked, me."); Ex. 182 at 8889 (April 2019: over 315,000 daily active people exposed to SSI content through algorithmic recommendations on Instagram Explore); Ex. 179 at 6750 (2020: "[with] the inquest into Molly Russell's death…the focus will be on what we were recommending Molly – and most crucially whether we have done enough to prevent 'future deaths.'….On search we're exposed with nowhere to hide."); Ex. 168A at 269:20-270:8. They were informed by an audit committee that "Instagram does not appear to adequately remove, label, or demote pro-restrictive eating disorder content, nor pro-suicide and/or pro-self-harm materials." Ex. 183 at 0074. They were even informed by parents (via a Meta-sponsored focus group) that they should "alert parents, schools, [and] emergency services" about this issue. Ex. 181 at 4759; *see id.* at 4758 ("It absolutely confounded them that we could be working on an issue so long, yet we are where we are.").

Characteristically, that is not what Meta chose to do. Instead, it maintained a "reactive" posture to removing SSI content. Ex. 182 at 8907 (3/4 of SSI content flagged reactively rather than proactively); Ex. 176 at 4022 (with respect to "viral suicide & self-injury challenges," "[w]e've historically taken a reactive

approach … (i.e. we deal with it once it's become so viral that media outlets catch wind of it."); Ex. 2A at 255:5-256:21. It rolled out new features on which SSI content couldn't be flagged, let alone removed. *Id.* at 48:15-49:16, 53:2-20 (explaining how classifiers were "very ineffective on short-form video," a problem that Meta "[w]as not able to fix before the launch of Reels."). And it published misleading statistics minimizing these problems. *Compare* Ex. 147 at 1 (public-facing "Community Standards Enforcement Report" estimating that "the upper limit" of content violating Meta's suicide and self-harm policy was "0.05%.") *with* Ex. 145 at 5049 (internal survey estimating 6.7% exposure to SSI). As one former employee acknowledged, Meta's public data presented a "deliberately incomplete picture in an underreporting of the rate of harmful content and experiences that were happening on our platform." Ex. 3 at 457:22-458:1; *see also* Ex. 2A at 222:6-12 (acknowledging disconnect between Meta's SSI policy and the aggregated content recommended by Instagram).

Most shockingly of all, Meta has lied about its approach to removing child pornography on its platforms. Since 2015, Instagram's Community Standards have stated, "We have zero tolerance when it comes to sharing sexual content involving minors or threatening to post intimate images of others." Ex. 184 at 11. This is unquestionably the right thing for Meta to say. But it is also unquestionably false. Meta operates three AI tools (or "classifiers") meant to detect content that likely violates its policies against child sexual abuse material ("CSAM"), CSAM solicitation, and child sexualization, respectively. Ex. 166 at 29-30. One would hope that, if these tools determine content more likely than not violates these policies, it would be automatically removed. Meta does not take such an approach—not even close. To the contrary, even if these classifiers determine *with 100% confidence* that photos or videos violate its policies against CSAM and child sexualization, Meta will *not* automatically delete that content. *Id.*; Ex. 985 ¶ 106. This was an intentional choice. In 2021, Meta product managers acknowledged they "could reduce [the] precision thresholds for quarantine or auto-delete" of content violating its child safety policies. Ex. 188 at 5478. (Translation: Meta could go from auto-deleting *none* of the content flagged by these classifiers to at least auto-deleting *some* of it.) But they decided against doing this, citing a "high risk" of false positives, "harsh penalties for child safety offenses," and a "[l]ack of appeals staffing (so inability to correct account-level disables"). Ex. 188 at 5478. It goes without saying that Meta could have solved this problem by simply hiring more appeals staff. Instead, it took the approach most likely to keep known child

pornography on its systems. This is the opposite of zero tolerance. *See also* Ex. 189 at 1907 & Ex. 2 at 371:13-373:20 (interstitial allowing users to "see results anyway" in response to search query that "may be associated with child sexual abuse").

Meta's human content moderation for child sexual abuse material has been just as deficient as its automated systems, and Meta has also lied about its efficacy. In March 2020, Mr. Zuckerberg told reporters that Meta was "actually surging the number of people who are working on" moderating "the most sensitive types of content," including "work on child exploitation." Ex. 187 at 3403. This was another lie. Ex. 2 at 357:11-23 (Q: "in fact, Meta was not surging the number of people working on child exploitation, correct?" A: "Not the total number of people, no." Q: "And as a consequence of that, the capacity to review potential child sexual abuse material on Instagram went down significantly during COVID; isn't that correct?" A: "Yes, that's correct."). Indeed, as of November 2020, Meta's "human review capacity" for child exploitation imagery was "at ~30%" of what it needed to be—with "<20% of projected supply for the near future." Ex. 187 at 3402 & 357:18-358:14; *see also* Ex. 186 at 9785 (2021: "Up until last half…no workstream set up to do bulk enforcement against FB Groups with _clearly_ violating content in either CEI [child exploitation imagery] or MS [minor sexualization]")

The consequences of Meta's upsettingly high tolerance for child pornography and sexualization have been devastating. *See* Ex. 2A at 272:3-9; Ex. 185 at 1113 (2020: "Is anyone else receiving child abuse reports on the platform? In the last 24 hours I have received a ton… I've seen terrible things, we're not detecting it."); Ex. 210 at 2309 ("21x fewer" deletions of child exploitation imagery on Instagram compared to Facebook), 2310 (over 10,000 users following known child exploitation related hashtags such as "#sexykid" and "#13yearoldboy"), 2311 (over 27,000 threads between bad actors and a teen over a 7 day period), 2314 (providing examples of "accounts sexualizing teens"). *See also id*. at 2324 & Ex. 2 at 302:3-303:20 (simply typing "h" into search bar could return results including "hottieskids" and "hot young boys"), *id*. at 2327 & Ex. 2A at 305:16-306:16 (searching for "childporn" returns numerous hashtag results with those words, including "#childporncollection" and "#childpornisgood"); Ex. 325 at 0324 & Ex. 2A at 286:16-25 (for test account created by Meta, "CEI being recommended as the top result in the Explore section").

The examples discussed above reflect the reality that child safety is simply not a priority at Meta—as additional internal documents make explicit. *See* Ex. 363 at 0155 ("Child Safety is *explicitly* called out as a non-goal in our H2 plans."); Ex. 378 at 9908 (saying that Meta's efforts on child grooming are "somewhere between zero and negligible" and adding "Child safety is an explicit non-goal this half"); Ex. 1068 at 5137-38 (showing there were five engineers, two data scientists, and one data engineer to cover "all child safety problems" including "the multiple other problems that [Meta] have to 'non goal' on – no sexual child abuse – child nudity").

At no point did Meta warn parents that it had only a fraction of the bodies it needed to review potential child sexual abuse material on Instagram. Ex. 2B at 358:6-21. Indeed, as recently as January 2024, Senator Josh Hawley asked Mr. Zuckerberg if Meta's platforms were "a disaster for teenagers" and Mr. Zuckerberg responded: "Senator, that's not true. That's not true." Ex. 467 at 37. But it is true. *The same day* Mr. Zuckerberg testified, Meta's employees were declaring a "SEV" (sitewide emergency event) related to child safety—after a team "discovered that a number of Child Safety banked search terms" were not being applied to block results on Instagram Search. Ex. 468 at 1412.

Meta has never publicly corrected its zero tolerance policy or otherwise warned school counselors, teachers, parents, or kids that its policy was not being enforced, manually or automatically. Ex. 2A at 313:21-25 (acknowledging inconsistency), Ex. 2A at 314:2-315:17, 474:18-22. Nor did Meta warn the public that its algorithm sometimes recommends child pornography to Instagram users. Ex. 2 at 262:10-263:4 (Q: "In your experience, did Instagram's algorithm, in fact, sometimes recommend to its user sexual content involving minors?" A: "Yes."), 286:16-25 (Instagram's internal investigation indicated that CSAM was being recommended in Explore). Instead, Meta continues to falsely maintain that "[o]ur systems are effective at reducing harmful content." Ex. 204 at 4 (quote in *Wall Street Journal*); *compare id. with* Ex. 2B at 381:23-382:3 (Q: "Based on your experience working at Meta for three and a half years on child safety issues, do you think it's accurate to say that Meta's systems are effective at reducing harmful content related to child sexualization and child exploitation?" A: "No."). Meta's willful and knowing disregard of the rights of others.

(7) Meta's willful and knowing disregard of the rights of others

Meta has a unique responsibility to protect the youngest users of its platforms, as its own

employees have acknowledged. Ex. 379 at 7816 ("we have more responsibility to protect teens"). Instead of embracing that responsibility, Meta has prioritized its own profits—over and over and over again. The preceding pages show how this issue played out, over years and across initiatives.

The Districts have submitted an expert report from a corporate ethicist explaining how Meta violated its own code of conduct and engaged in "pseudo corporate social responsibility." *See* Ex. 993 ¶¶ 156-388. But testimony from Meta's own employees says it all. Vaishnavi Jayakumar, former Head of Safety and Well-Being at Instagram from 2020-2023, testified that it was "generally pretty difficult to make safety changes that might impact growth or daily active users by any significant amount." Ex. 2B at 384:22-385:2. Dr. Joshua Simons MP, a research scientist at Facebook AI from 2018-2022, testified that, "if we were proposing something that would reduce engagement that went up for the executive team, Mark Zuckerberg, to review, the decision that came back would prioritize the existing system of engagement over other safety considerations." Ex. 13 at 60:15-61:11. [redacted] a senior UX researcher at Meta from 2017-2022, regretted that she "joined the company a little naïve about what the company would want to do" and that research-backed recommendations "were rejected because the changes would have limited certain metrics, like engagement." Ex. 10 at 34:19-23. Brian Boland, whose eleven-year tenure at Meta ended as VP of Partnerships, testified that, "My feeling then and my feeling now is that they don't meaningfully care about user safety. It's not something that they spend a lot of time on. It's not something that they think about. And I really think they don't care." Ex. 8 at 109:2-16.

These are individuals from different parts of the organization, who worked there at different times, with varying degrees of seniority. Their perspectives are entirely consistent. And they are not alone. *See also, e.g.*, Ex. 3 at 460:21-461:5 (Allison Lee: "they perhaps cared as a third or fourth order priority about safety and well-being"); Ex. 23 at 247:7-10 (Frances Haugen: "ongoing trend of not prioritizing safety as high as it should have been"); Ex. 168B at 482:20-21 (Arturo Béjar: "prioritize[d] engagement and growth over safety and well-being."); Ex. 202 at 1 (Sarah Wynn-Williams: leadership "maintained their focus on growth" even though "[t]hey knew the negative externalities of their products were being pushed onto teens"); Ex. 256 at 59:8-20 (Lotte Rubaek: child safety "not at all" a priority).

Despite earning billions of dollars in annual revenue—and its leader being one of the richest people in the world—Meta simply refused to invest resources in keeping kids safe. One final story sums

this up. In April 2017, Mr. Systrom wrote to Mr. Zuckerberg appealing for additional staff to address a variety of safety issues as to which "women and teens [are] particularly impacted." Ex. 663 at 6745. Mr. Systrom emphasized the importance of upholding his "public commitment to making Instagram a place where people feel safe." Ex. 663 at 6744. Mr. Zuckerberg responded: "we're facing more extreme issues on FB right now….I do view funding that as more urgent in the near term." *Id*. He continued, "I probably can't get you 13 engineers in the near term, even though I'm supportive of these projects." *Id.* Mr. Systrom groused to his team: "I don't think Mark understands the urgency of working on integrity related issues at IG." *Id.* at 6751.

By October 2017, the situation had not improved. According to an internal quarterly review, Instagram's protect and care (PAC) team remained "far behind" that of Facebook—"we could become a major liability for FB Inc." Ex. 666 at 1333. "Instagram is now over 850M monthly users, and with that size comes the responsibility to ensure our community is safe. Also, we're launching more new products at a faster rate than ever, each of which require additional work to prevent abuse and reduce exposure of FB Inc. We're continuing to see an increase in high intensity abuse." *Id*. at 1334. "However, the PAC team is only 22 engineers. We aren't staffed …to stay ahead of all potential PAC related issues." *Id*.

Fast forward two years. On April 8, 2019, a senior executive at Meta wrote to Mr. Zuckerberg informing him that the company had a "deep understanding around three negative drivers that occur frequently on FB and impact people's well-being," including "problematic use (prevalence: 55% mild, 3.1% severe)" and "social comparison (prevalence: 40% mild, 5% severe)." Ex. 398 at 5473. The executive asked Mr. Zuckerberg for an additional headcount of 24 heads, to fund a "Well-being 10x" initiative. *Id.* "We believe there is a strong need to increase our investment in these areas to make a meaningful shift in these areas." *Id.* Mr. Zuckerberg did not respond. But Adam Mosseri did: "I don't see us funding this from Instagram any time soon." *Id.* "The 24 head count were not funded" and the team was asked "to do the work with existing head count that they have." Ex. 399 at 226:15-18.

Fast forward another two years. On August 27, 2021, Meta's Head of Global Affairs, Nick Clegg, wrote to Mark Zuckerberg with a "proposal…for additional investment to strengthen our position on wellbeing across the company." Ex. 278 at 0952. "Our wellbeing work is both under-staffed and fragmented across teams," he wrote. "We are not on track to succeed for our core well-being topics,"

including "problematic use … and SSI." *Id.* "[I]f not addressed, these [problems] will follow us into the Metaverse." Ex. 278 at 0952. Again, Mr. Zuckerberg did not reply.

Two months later, in October 2021, Arturo Béjar wrote to Mr. Zuckerberg. Mr. Béjar had served as the head of Facebook's Protect and Care team for years and was then working as a consultant for Meta. He warned Mr. Zuckerberg that, per internal research, "51% of Instagram users say 'yes' to having a bad or harmful experience in the last 7 days" and nearly "39.4% of 13-15 year olds said they experienced negative [social] comparison." Ex. 303 at 4804-05. Mr. Béjar went on to identify well-being solutions and tell Mr. Zuckerberg a variant of what other executives had been saying for years: "I believe that it is important to get the[se] efforts well-funded and prioritized." *Id.* at 4805. Mr. Zuckerberg did not reply.

That same week, Frances Haugen publicly testified before Congress about hundreds of internal research documents that Meta had kept secret for more than a decade, many concerning the impact of Instagram and Facebook on youth mental health. Meta responded by sending Mr. Mosseri to the Hill to placate Congress. Mr. Mosseri testified: "As Head of Instagram, I am especially focused on the safety of the youngest people who use our services." Ex. 164 at 2. But behind closed doors, Mr. Zuckerberg was texting confidants, "I'm not going to say [child safety] is my personal main focus when I have a number of other areas I'm more focused on like building the metaverse." Ex. 664 at 2509.

In November 2021, Nick Clegg renewed his efforts to acquire additional funding for well-being work, sending a "revised investment proposal" that "scal[ed] down" his initial funding request. Ex. 278 at 0951. "Given the increased urgency of all this since the initial escalation in August, are you supportive of funding this scaled back proposal?" he asked. *Id.* Despite this plea, no additional headcount was forthcoming. Ex. 278 at 0950; Ex. 9B at 472:16-17. Mr. Zuckerberg's response? Silence.

b) TikTok's Conduct

(1) TikTok's business model

TikTok, like its co-defendants, built its extraordinary success on one commodity: attention. Its business model depends on capturing and holding users—especially school-aged children—for as long as possible, because every additional minute on the platform translates directly into revenue. Together with its parent company, ByteDance, TikTok monetizes that attention through what its own employees described as an "advertising business model" that "create[s] incentives for companies to grow at the

expense of its users." Ex. 632A at 50:5-10, 103:16-17. Executives have acknowledged that advertising provides the vast majority of TikTok's revenue. Ex. 636 at 47:11-21. Internal documents confirm the priority: [redacted], and [redacted], set out to "[a]ccelerate ads revenue to [redacted]" in a single quarter of 2023—up from 2022's [redacted] per day, or roughly [redacted] annually. [redacted] at 47:3-10; Ex. 522 at 6223; [redacted] at 451:1-6; Ex. 523.

[redacted] declared the corporate governing principle that would drive TikTok's every decision: "With more active users, more clicks, and better user satisfaction comes higher traffic, which is then turned into revenue." Ex. 434 at 7156; [redacted] at 115:7-117:15. That was not an aside—it was a directive. Inside TikTok, the connection between user engagement and profit was not merely understood; it was institutionalized. An internal document—confirmed by TikTok's former [redacted]—put it bluntly: "the advertising-based business model encourages optimization for time spent in the app." Ex. 541 at 9176; [redacted] at 60:15-61:1. The logic was as simple as it was deliberate: more time meant more ads, and more ads meant more money—billions!

TikTok pursued that attention-for-profit model relentlessly. Internal documents show that in just one quarter of 2023, the company targeted [redacted] daily active U.S. users. Ex. 477 at 44:20-45:16; Ex. 522 at 6223. "Stay duration"—the time users spent in the app—was treated as a "top-level metric." Ex. 635 at 410:15-19; Ex. 525. The mandate was clear: increase average use from [redacted] per day in 2021 to [redacted] in 2022. Ex. 635 at 455:9-23; Ex. 526. To achieve that goal, TikTok deployed what its own records describe as "powerful coercive design tactics" that "tend to benefit companies and advertisers more than users." Ex. 632A at 103:20–104:10; Ex. 541 at 9185. These were not accidents of user behavior—they were deliberate design choices engineered to maximize time, attention, and revenue.

Even TikTok's supposed "safety" features were judged by a single yardstick: whether they hurt engagement—not whether they protected kids. When TikTok tested weekly screen-time updates for minors, the internal takeaway was that the feature had "neutral" effects—no change in daily active users, stay duration, or retention. Ex. 405A at 199:13–200:5; Ex. 488 at 7711. Another internal report, tracking a "Screentime Management Upsell for Minors," listed "Stay Duration" as a "Guardrail Metric" and

proudly noted "no statistically significant change." Ex. 405A at 203:6-204:11; Ex. 487 at 7234. The irony could not be clearer: these so-called guardrails were never meant to protect children—they were meant to provide the appearance of safety while ensuring that neither engagement nor revenue were ever meaningfully reduced. In short, "safety" was acceptable only if it left the bottom line untouched. When one test showed a [redacted] decline in stay duration, the report reassured executives that the drop "align[ed] within guidance" and had "no impact to retention." Ex. 635 at 431:3-15; Ex. 497 at 2841. That same report was blunter still, acknowledging "minors do not have executive mental function to control their screen time." Ex. 497 at 2841.

The picture could not be clearer. TikTok's business model does not tolerate limits on user attention—it thrives on eradicating them, even at the expense of safety. Every design choice, platform tweak, and so-called "safety" feature was measured by one metric: did it keep users, especially children, on the platform longer? The longer the stay, the more ads TikTok served, and the more revenue flowed to ByteDance. That was no accident. It was *the* overarching strategy—engineered, enforced, and embedded in the corporate culture. What mattered most was not safety, not user well-being, but screen time.

(2) TikTok's targeting of school-aged children

From the start, TikTok identified school-aged children—particularly middle and high school students—as its core audience. In 2018, after ByteDance acquired Musical.ly and rebranded it as TikTok, internal documents show that ByteDance and TikTok knew a large share of their users were "middle school and high school females." Ex. 411 at 63:2-16, 65:12-66:5, 66:10-66:21; Ex. 412 at 5197. Rather than distancing themselves from this young audience, they embraced them—formally defining school-aged children as the platform's "core audience." [redacted] at 116:18–25; Ex. 414 at 0624 (former [redacted] admitting that "high schooler[s] is a core audience of our platform"). TikTok made clear it wanted school-aged children to be its heaviest users, announcing a "rebrand" aimed at ensuring those ages 13 to 18 would be "heavy user[s]." Ex. 415 at 0205. Internal planning documents went further, identifying "high school boys" as a "[k]ey population group[]." Ex. 422 at 0220A; Ex. 413B at 109:14–20.

To better exploit this demographic, TikTok launched an internal initiative—"*Project Cool Kid*"—to study and capture the habits of school-aged users. The project focused on "how High Schoolers, predominantly L2 [15-17-year-old] males are using TikTok," to "understand high school trends," and to

learn what "fuels usage and adoption in High Schools." Ex. 423 at 7326; Ex. 411 at 132:6-134:11. TikTok and ByteDance wanted to know "how mainstream TikTok [was] in high school and whether or not all high schoolers are using TT." Ex. 424 at 9792; Ex. 413B at 127:18–24.

As part of this research, TikTok conducted "street interviews" with middle and high school students. Ex. 411 at 166:17-168:12. One [redacted], reported: "dude I talked to around 30 high schoolers," and boasted, "literally all middle school to high schoolers, like 8–18, have heard of TikTok in Houston and Chicago." [redacted] at 164:1-12, 176:10-13; Ex. 425 at 2374; Ex. 426 at 5153. These interviews were not confined to junior staff. Top executives from both ByteDance and TikTok—including [redacted], and [redacted]—personally walked the streets conducting interviews with school-aged children. [redacted] at 165:10-166:14. TikTok conducted these interviews annually. [redacted] at 184:19-21. The findings were telling: children as young as eight used TikTok "every day," while others admitted that "everyone says you'll get addicted." Ex. 411 at 186:10-15, 211:5-22; Ex. 428 at 5042; Ex. 429 at 14.

Armed with this data, TikTok set out to capture the group of adolescents it had a thin hold on from Musical.ly—high school males. It launched targeted advertising campaigns on co-defendant Snap's platform, Snapchat "with the aim to grow L2 [high school-aged, 15-17-year-old] male users," shifting ad spend from Facebook to reach a younger audience. Ex. 412 at 5220; Ex. 411 at 109:22-110:23. TikTok quickly tripled its daily ad spending to approximately [redacted] Ex. 508 at 269:24-270:11; Ex. 430 at 3467. They even removed "age filters on paid [ad] channels in the U.S. market" to maximize reach among teenage males. Ex. 508 at 137:2-139:3; Ex. 431 at 8836. These efforts were central to TikTok's 2019 objectives and key results: to "attract teen male users" and "drive market growth." Ex. 413B at 111:15-23; Ex. 422 at 0220A.

The strategy worked. By the second half of 2019, the percentage of 15–17-year-old males among daily new users "grew significantly...reaching [redacted] and surpassing all other user groups except for L2 [high school] female[s]." Ex. 411 at 113:21–114:3; Ex. 412 at 5221. Buoyed by this success, TikTok declared itself "mainstream" in high schools. Ex. 411 at 124:19-125:6; Ex. 412 at 5196, 5220-5221. Even as it later sought to "age up" its user base, TikTok acknowledged it was "maxed out" with middle and high school users—and committed internally to not "eliminating or overlooking" users ages 13–17, the

core of its user base. Ex. 411 at 200:7-21, 373:25–374:5; Ex. 432 at 7167; Ex. 433 at 156:3-7.

In September 2021, TikTok celebrated its own success with a video press release titled "*Thanks a Billion*," boasting that "more than 1 billion people from around the world now come to TikTok every month." The announcement featured [redacted], the same executive who had walked the streets of American cities interviewing students. *See* Ex. [redacted] at 135:13-20, 136:3-137:20; Ex. 569.

TikTok has been staggeringly successful at attracting young people to its platform. In sworn discovery responses, TikTok confirmed that users ages 14–18 skyrocketed from [redacted] in 2018 to more than [redacted] by 2022 [redacted] in 2024:

Year	Users ages 14-18[6]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

Ex. 410 at 22–23. These figures confirm that TikTok's campaign successfully entrenched its dominance among middle and high school students—making the resulting harms to schools not only foreseeable but inevitable. By June 2022, TikTok had achieved a [redacted] **penetration rate among children aged 13–17**. Ex. 433 at 141:1-6; Ex. 538 at 6. As [redacted], admitted, a [redacted] penetration rate was [redacted]. [redacted] at 151:13-152:3.

With that dominance came advertisers—and money. ByteDance's [redacted], made TikTok's priorities clear: to acquire and retain users while "[g]enerating revenue." [redacted] at 118:18-119:7; Ex. 434 at 7155-56. TikTok's year-end business reviews echoed that mandate. The 2021 review set a target of [redacted] billion in advertising revenue and noted that average daily time spent on the app had climbed to [redacted] minutes. Ex. 437 at 134:5-9; Ex. 435 at 7029. Internal records confirm the underlying philosophy: an "advertising-based business model [that] encourages optimization for time spent in the app." Ex. 436 at 9176.

[6] TikTok did not provide statistics that included 13-year-olds, omitting a significant portion of its underage user base.

To sustain its growth, TikTok marketed its young audience directly to advertisers. [redacted] at 203:3–24 ([redacted], confirming that "Gen Z was a very—a very common topic of conversation with advertisers"); Ex. 438 at 3, 5 (noting [redacted] of users are Gen Z and that "Gen Z watch and interact with more videos per day"). In 2019, TikTok's brand partnership deck touted explosive U.S. user growth and described the platform as "Generation Z Focused." Ex. 437 at 99:25-100:2; Ex. 439 at slide 14. The same presentation emphasized that users opened TikTok more than [redacted] daily and spent over [redacted] a day on the app. Ex. 437 at 103:1-7; Ex. 439 at slide 15.

TikTok made similar appeals to individual brands. In a presentation to [redacted], TikTok reported that [redacted] of its U.S. audience was between ages 13 and 24, highlighting "sustained engagement" and admitting that "our algorithm makes [the] app very addicting." Ex. 437 at 212:22-213:5; Ex. 440 at 9376, 78-79. In a pitch to [redacted], TikTok boasted that [redacted] of its audience was between ages 13 and 17—[redacted]—and that this cohort spent the most time on the platform, averaging more than [redacted] per day. Ex. 437 at 224:8-10; Ex. 441 at 9049–9050. TikTok also emphasized that Gen Z users were "[redacted]," reporting that [redacted] of its monthly active users were ages 13-17 as of April 2021. Ex. 437 at 230:10-231:5; Ex. 442 at 19.

The record leaves no doubt: TikTok deliberately targeted school-aged children, studied their habits, and built its platform around keeping them engaged. Its own data, internal communications, and marketing materials confirm this was not incidental—it was the primary business strategy.

(3) TikTok's targeting of schools

To cement its dominance among school-aged children, TikTok turned its attention to schools themselves—viewing them as the most direct path to widespread adoption. It sought to embed TikTok in the educational ecosystem, partnering with national organizations and developing initiatives that reached directly into classrooms. Like Meta, Snap, and Google/YouTube, TikTok borrowed from Big Tobacco's playbook—using "youth prevention" and school programs to infiltrate classrooms and burnish its image under the guise of education. Every strategy served a single goal: making TikTok an unavoidable part of students' daily lives.

TikTok began by disguising commercial objectives behind partnerships that appeared educational or community-driven. Chief among these was its collaboration with the National Parent Teacher

Association ("National PTA"), a respected organization representing millions of students, parents, and educators. In 2019, TikTok initiated what it described as a "formal relationship" with the National PTA, leveraging its reputation to gain access and legitimacy within schools. Ex. 413B at 182:7-183:11.

Internally, TikTok described the PTA partnership as an opportunity to "leverage" its vast network—54 state PTAs, 8,000 council PTAs, and 24,000 local PTAs, representing 4 million members and 16.2 million families. Ex. 443 at 4129, 4133. TikTok identified several advantages: (1) the PTA could "positively raise ByteDance's profile among parents"; (2) co-branded materials could reach "PTA leaders and influential, engaged families"; (3) TikTok could "tap[] into key National PTA campaigns and events"; and (4) the company could "draw[] on the community trust, 120-year-old established reputation, and perceived values of PTA." *Id.* at 4130.

Ultimately, TikTok saw the PTA partnership as a vehicle for influence. After securing a paid sponsorship, company officials noted the PTA would "do whatever we want going forward in the fall . . . [t]hey'll announce things publicly[,] [t]heir CEO will do press statements for us. They'll co-brand ConnectSafely's '*How to TikTok Guide*.'" Ex. 444A at 50:23-52:4; Ex. 445 at 4722. In October 2019, TikTok executed a Sponsorship Agreement with the National PTA [redacted]." Ex. 444A at 58:16-24; Ex. 446 at 8241. This clause [redacted], including a [redacted] in 2020 and an additional [redacted] in 2024. Ex. 444A at 60:15-61:4, 90:19-94:23; Ex. 447 at 0392; Ex. 448 at 4480.

TikTok then used the National PTA as a front to promote its "safety" initiatives, such as "Safer Internet Day." Ex. 444A at 57:4-58:2; Ex. 446 at 8239. It distributed "mass mailer" materials to parents, educators, and school districts promoting TikTok's purported safety features, including the *TikTok Guide for Parents* and *TikTok Family Pairing Guide for Parents.* Ex. 444B at 479:24-480:5; Ex. 449; Ex. 450. Internally, it celebrated that "tons of schools" printed the guide and left copies "in the principal's office." Ex. 451 at 4792.

These partnerships were not public service—they were public relations. By laundering TikTok's image through a trusted national institution, TikTok sought to reassure parents and educators, deflect scrutiny, and preserve the engagement driving youth harm.

To deepen its reach, TikTok co-hosted "flagship" school events with the National PTA nationwide. Ex. 413B at 162:3-13, 182:9-19; Ex. 448 at 4493. One such event at Westwood High School in November 2022 drew 300 parents, students, and administrators. Ex. 452 at 3744. Despite the family-friendly veneer, these events revealed TikTok's knowledge that children under 13 were active users. Ex. 444A at 222:16-228:5; Ex. 453 at 4466 ("This student panel is primarily under 13").

At another event in February 2020 at a Los Angeles school, a student panel featured children under 13 who admitted they "lie[d] about their birthdate to get around" TikTok's age restrictions. Ex. 444A at 225:7-13, 226:10-20; Ex. 453 at 4466. Internally, executives noted they "seriously dodged a bullet" when CBS and ABC canceled planned national coverage. Ex. 444A at 225:23-226:3; Ex. 453 at 4466.

A central goal of the partnership was favorable publicity. TikTok's sponsorship agreements required that TikTok appear on the National PTA's "sponsors and partners" webpage with a link to ByteDance's site and mandated that the PTA provide "positive quotes" about TikTok's sponsorship. Ex. 446 at 8241; Ex. 448 at 4495.

TikTok capitalized on these commitments immediately. In 2020, TikTok launched a "Trust & Safety" marketing campaign across twelve politically strategic states—chosen for overlap with key congressional committees—and used the PTA to bolster its credibility. Ex. 454 at 7940. The PTA also helped promote TikTok's "Family Pairing" feature, publishing op-eds offering "commentary in support of TikTok's tools and resources," which TikTok called "of enormous value." Ex. 455 at 6205; Ex. 444A at 63:5-64:9. TikTok also used the PTA partnership for political influence. TikTok targeted event sites in "major market media centers" and "sensitive political districts," noting that "Tampa gets us both Senators Rubio and Scott." Ex. 456 at 9201, 9202. Internally, it described National PTA grants funded by TikTok as "[m]ore good news to share with policymakers!" Ex. 457 at 8844.

Across all channels, TikTok controlled the PTA's messaging. TikTok employees directly edited PTA quotes and supplied talking points for use with local leaders. Ex. 458 at 6228; Ex. 459 at 2229. The PTA even agreed not to participate in interviews "without first having a planning session" with TikTok's media team. Ex. 460 at 1997. To enforce this control, [redacted]. Ex. 444A at 158:14-159:5; Ex. 446 at 8237.

Beyond the PTA, TikTok paid other educational and "digital safety" organizations to promote its

image. It paid Common Sense Networks [redacted] "to deliver an education and awareness program" that presented TikTok as empowering teens to "take ownership of its digital lives." Ex. 444A at 115:18-117:5; Ex. 461 at 3873. TikTok cited this partnership in its *TikTok Guide for Parents,* distributed nationwide through the National PTA. Ex. 444A at 284:16-25; Ex. 449 at 6392.

TikTok also paid ConnectSafely, a nonprofit, to "draft and publish a 'Parent's Guide to TikTok'" and co-host events like Safer Internet Day. Ex. 444A at 119:5-16; Ex. 462 at 4071. Internally, TikTok described ConnectSafely's CEO, Larry Magid, as "a good shield for us going forward," able to "serve as a public advocate for [TikTok] in the face of all the criticism." Ex. 444A at 246:7–247:2; Ex. 463 at 1110. TikTok publicly cited the partnership—including in CEO Shou Chew's Senate testimony—[redacted]. Ex. 444B at 349:04-350:14; Ex. 464 at 3837; Ex. 465 at 2626.

Taken together, TikTok's network of sponsored partnerships—with the National PTA, Common Sense Networks, and ConnectSafely—was not a safety initiative but a coordinated public-relations campaign. Like Big Tobacco and JUUL, TikTok repackaged corporate promotion as public education, exploiting schools' and parents' trust to entrench its platform among youth. *See* Ex. 662 (2019 congressional hearing regarding JUUL's role in youth nicotine epidemic).

TikTok also developed direct plans to enter classrooms, aiming to make TikTok "the #1 short video app for creation and consumption of educational content" and a "resource for classrooms" where teachers and students could "build curricular content on TikTok at scale." Ex. 469 at 0563; Ex. 470 at 9753; Ex. 444A at 126:3-25. It envisioned "school-specific Discover Pages" where educators could post content and assign "homework" as TikTok videos, to be piloted in major school districts. Ex. 471 at 6534.

To promote this initiative, [redacted]. Ex. 471 at 6529; Ex. 472 at 9074. These influencers—branded as "Educreators"—were presented as "standout teachers from the TikTok community" [redacted] to encourage others to post under #TeachersOnTikTok. Ex. 474 at 3472. TikTok sought media coverage of these campaigns, coordinating with the National PTA to amplify hashtags such as #LearnOnTikTok and #TeachersOnTikTok. Ex. 475 at 6980, 6987; Ex. 476 at 8578. Internally, it credited these efforts with boosting engagement under the #BackToSchool tag. Ex. 473.

Finally, TikTok leveraged its massive school-aged audience to attract advertisers during back-to-school season. Once it secured a critical mass of student users, it promoted TikTok as the premier platform for "back to school" campaigns—boasting an audience "full of school-aged children." Ex. 437 at 245:11-246:2; Ex. 473. It told advertisers to expect "an influx of school supplies and clothing hauls from students and parents alike," and "endless opportunities for brands to join in on the playground fun." *Id.*

TikTok closely tracked hashtags like #BackToSchool, touting a "massive jump in viewership" from its "diverse audience of Gen-Z students, edu-Creators, and parents." Ex. 437 at 246:8-16; Ex. 473. It knew children were using TikTok in school: "Between classes, during lunch, and after that final bell rings, Gen-Z users... are pulling the squad together to dance, participate in trends, and create on TikTok." Ex. 437 at 247:12-25. It even instructed advertisers how to appeal to this demographic—targeting #collegeLife, #teacherLife, and #highSchool communities, promising, "you'll be the coolest kid in school in no time." Ex. 437 at 248:20-249:7; Ex. 473.

(4) TikTok's knowledge of harms

Despite its heavily promoted platitudes about "digital safety," TikTok knew the design of its platform posed serious risks—especially to minors. In a 2020 internal safety review, [redacted], stated that social media is "inherently dangerous" and that minors are "at risk" on TikTok. Ex. 530 at 4753; [redacted] at 253:12-254:2. Through commissioned research, external partnerships, user feedback, and internal studies, TikTok understood how its platform caused or worsened addiction, anxiety, depression, sleep disruption, poor body image, and suicidality—and that it disrupted students and schools.

(a) TikTok's addiction problem was known, acknowledged, and profitable

Shortly after TikTok launched in the United States, TikTok knew its platform was habit-forming and that it was fueling addictive and problematic use among children. As early as March 2019—barely months after ByteDance rebranded Musical.ly as TikTok—executives received reports directly from parents warning that "kids have become addicted to TikTok." Ex. 532 at 7593. One internal "Product Experience Report" captured TikTok's attitude succinctly: while parent concerns about youth addiction were "concerning for young minds," this phenomenon was viewed internally as "a good thing for the platform." *Id.;* Ex. 635 at 360:4-19. Addiction, in other words, was not an unforeseen byproduct—it was

treated as a performance indicator.

The complaints multiplied year after year. Between 2018 and 2024, TikTok received countless reports from its users and parents describing TikTok as "addictive," "like crystal meth," and "ruining" its children's attention and sleep. Exs. 583, 584, 585, 586, 587, 588, 589, 590, 591. Users told TikTok they were deleting their accounts because they were "getting addicted" or "trying to get over [its] TikTok addiction." *Id.* Teachers and parents reported that students were "getting bad reports from teachers" because of TikTok use. Ex. 589 at 2.

Internally, these warnings were heard but not acted upon. By 2020, TikTok's own internal documents are replete with admissions like "addiction to technology is a ubiquitous problem that TikTok and most other platforms deal with today." Ex. 534 at 1694; Ex. 632A at 228:8-21. The phrase "ubiquitous problem" was not an excuse but an admission: TikTok recognized that its platform was driving the same addiction pattern across millions of users, particularly school-aged users.

Even external consultants hired by TikTok confirmed this reality. [redacted], a behavioral research group retained by TikTok, warned that the addictive experience was not incidental, it was intentionally engineered. Ex. 581 at 0111. Yet TikTok continued to design and promote engagement-maximizing features known to drive compulsive use, choosing profit over prevention.

The result was an established pattern of knowledge and inaction. Year-after-year, internal reports, user feedback, and expert findings documented the same conclusion: TikTok was fostering addiction among its users, particularly minors. TikTok knew it, quantified it, and continued to capitalize on it.

(b) TikTok's executives and employees struggled with addiction and its co-Defendants knew how addictive TikTok was

Even TikTok's own executives and employees struggled to control their use of the platform. Their personal experiences confirmed what TikTok's internal data had long shown: the product was engineered for compulsion. [redacted], admitted under oath that he deleted the app and installed a third-party blocker, the Freedom app, to curb his own use. [redacted] at 277:1-278:9. [redacted] was not a teenager whose brain was still developing—he was a senior safety executive who needed outside software to stop himself from using the very platform he was responsible for making safe. *Id.* at 281:24-

282:4. His experience laid bare the depth of TikTok's design: it could hook even the people charged with protecting children from it.

[redacted] was not alone. [redacted], admitted that "TikTok is quite addicting." Ex. 676 at 3304. [redacted], reached the same conclusion within two months of joining the company: "I'm addicted." Ex. 681 at 6089. These were not casual users—they were the architects and guardians of the product, each personally experiencing the dependency their platform produced in millions of children.

The recognition of TikTok's addictiveness extended beyond its own walls. Google—one of TikTok's co-defendants and fiercest competitors—internally concluded that "TikTok can be extremely addictive." Ex. 794 at slide 27. Similarly, [redacted], who testified in this action as Meta's corporate designee, privately acknowledged to a colleague about TikTok: "I deleted the app because I was … addicted." Ex. 1180 at 3628 (ellipsis in original). These are not outsider's critiques but acknowledgments from rivals that fully understood the mechanics of engineered engagement. These assessments confirmed what TikTok's insiders already knew: the platform was deliberately built to keep users hooked.

These admissions carry unique weight. [redacted] were insiders with direct access to TikTok's engagement metrics and safety systems. Their struggles were not coincidences—they were symptoms of a product designed to capture attention and resist restraint. If senior executives—experts in the very features driving addiction—could not limit their own use, the impact on children, whose self-control and judgment are still developing, was not only foreseeable but inevitable. TikTok's own leaders, not to mention its rivals, recognized what TikTok now pretends not to— addiction was not a side effect of its design, but its success metric.

(c) TikTok intentionally engineered addiction

TikTok's own internal data and research confirmed what users, parents, and external experts had long been saying: TikTok's platform was compulsive and habit-forming, producing patterns of excessive use that harmed mental health and well-being. Far from being unaware, TikTok rigorously studied these effects through surveys, focus groups, and internal analytics—and repeatedly documented addiction-related concerns in its internal files. To date, TikTok has hidden those files from the public.

As early as 2020, TikTok's internal surveys showed that a significant share of users actively tried to reduce or quit using TikTok because they were spending "too much time" on the app. Ex. 535 at 2294. In these surveys, "fear of getting addicted" ranked among the leading reasons users deleted the app altogether. Ex. 536 at 3102. Users repeatedly described feeling hooked and unable to stop scrolling, expressing guilt and frustration over the hours lost each day on the app. Ex. 537 at 3040. These findings were strikingly consistent across demographic groups, but particularly pronounced among youth, underscoring that the problem was not isolated—it was systemic.

A 2021 internal analysis by TikTok's Trust & Safety team went even further. It explicitly acknowledged that "our users' biggest usage deterrent is that they think the platform is addictive," and that "compulsive usage correlates with a slew of negative mental effects, like loss of analytical skills, memory formation, contextual thinking, conversational depth, empathy and increased anxiety." Ex. 537 at 3040-3041; Ex. 632A at 259:1-7. The same document warned that users were caught in a loop of automatic engagement, which deprioritized self-control. Ex. 537 at 3044. The internal analysis concluded candidly that "compulsive usage on TikTok is rampant and our users need better tools to understand its usage [and] manage it effectively." Ex. 537 at 3041; *see id.* at 3040 ("users' biggest usage deterrent is that they think the platform is addictive."); Ex. 632A at 257:13-25, 260:8-19.

TikTok's internal documents also make clear that the platform's addictive pull was engineered, not accidental. TikTok's behavioral research confirmed that the very features driving engagement were producing the cognitive dependence users described as "addiction." Ex. 537 at 3040-3041. Researchers noted that the platform's variable reward structure conditioned users to seek intermittent dopamine rewards through endless swiping, mirroring behavioral reinforcement patterns found in gambling. Ex. 508 at 179:20-25, 181:15-182:24. Internal documents described the "habit moment" —the point at which use transitions from voluntary to automatic—which occurred after approximately 260 videos in the first week. Ex. 538 at 29; Ex. 433 at 159:6-16. [redacted], confirmed that the platform intentionally included a "slot-machine effect" that drove repeated engagement and stripped users of meaningful control. [redacted] at 82:23-83:8; Ex. 541 at 9176; Ex. 542 at 1765. This recognition of the "slot machine effect" appeared frequently across TikTok's own internal documents. *Id.*; *see also* Ex. 556 at 1514.

These findings did not come from isolated teams—they reflected company-wide recognition of the problem. Product designers, data scientists, and Trust & Safety personnel all contributed to the analyses. Further, TikTok saw that these compulsive patterns carried measurable harmful psychological consequences. Internal data linked excessive TikTok use to heightened anxiety, sleep deprivation, and decreased academic focus—effects TikTok categorized as "negative mental effects." Ex. 537 at 3040-3041. [redacted] and others acknowledged that TikTok's optimization for engagement produced rampant compulsive usage, correlating with anxiety, sleep disruption, and loss of concentration in school. [redacted] at 229:7-12, 260:8-19. TikTok's own data confirmed that school-aged users were among the most affected. Internal analyses acknowledged that students were active on the platform late at night, arriving at school sleep-deprived, and checking notifications during class—behaviors the company recognized as foreseeable outcomes of its engagement strategy. Ex. 413B at 202:14-21; Ex. 643 at 8755.

TikTok could have addressed the fundamental issue: Its platform design was creating addiction and mental health risks. Instead, TikTok continued to optimize for "time spent" and "retention," not user health Ex. 556 at 1514. Internal documents make clear that TikTok's financial goals are directly tied to its uniquely "powerful" and "coercive" nature. *Id.* TikTok's internal dashboards [redacted]. Ex. 632A at 257:13-25. [redacted]. Ex. 556 at 1514.

In short, TikTok's own research (still not public) confirmed that addiction was not an unintended byproduct of the platform but was an integral, measurable outcome. The very executives and employees responsible for mitigating harm could not themselves disengage, underscoring the depth of the problem and the deliberate nature of the addiction. Rather than putting up guardrails to protect school-aged children, TikTok built guardrails only to protect its revenue—accepting widespread addiction as the price of maximizing engagement and, by extension, profit. Rather than mitigating these harms, TikTok reinforced them, because every additional minute on the platform translated into more advertising revenue. Rather than warning users, parents, and educators about the problem, TikTok buried its research.

(d) TikTok knew its platform was harming users' mental health, sleep, and schools - but chose profit over prevention

TikTok's internal documents leave no doubt: it knew its platform was harming users' mental health, disrupting sleep, and undermining the classroom environment. Those harms were not theoretical. Internal studies, surveys, user feedback, and external research commissioned by TikTok repeatedly identified that excessive TikTok use is correlated with anxiety, depression, and diminished self-esteem. *See, e.g.*, Ex. 594 at 6419 (Trust & Safety team conceding that "compulsive usage correlates with a slew of negative mental effects, including increased anxiety," and that the platform "exacerbate[es] preexisting mental health conditions such as depression and anxiety."); Ex. 632A at 259:1-7.

By 2020 at the latest, internal safety reviews warned that social media is "inherently dangerous" and that minors were "at risk" on TikTok. Ex. 530 at 4753; Ex. 402A at 253:12-254:2. An internal 2021 report captured the problem starkly: users described feeling "sucked in" and reported a range of negative emotions after extended use, including "guilt, frustration, irritation, disappointment, and anxiety." Ex. 592 at 2102. TikTok conducted numerous focus groups, which confirmed that teens "acknowledge that they sometimes lack agency on TikTok," and that this loss of control was directly tied to stress and anxiety. Ex. 593 at 9673.

TikTok also tracked and reviewed external studies that reached the same conclusion. Internal memoranda cited peer-reviewed research showing that minors who spend more than three hours a day on social media double their risk of "poor mental health outcomes, including symptoms of depression and anxiety." Ex. 595 at 0819; Ex. 596 at 0680. Another study, included in TikTok's *Minor Safety Strategy Paper*, linked six or more hours of daily screen use to heightened depression risk. Ex. 485 at 0816. Yet TikTok's own data confirmed that its users—especially minors—were routinely exceeding that threshold. Nearly [redacted] users spent more than [redacted] on TikTok, prompting an internal acknowledgment that a "stronger approach" was needed to curb "harmful screen time." *Id.* [redacted], personally recognized that "depression and addiction" were "conditions often linked to the use of apps and social networks," acknowledging that these risks gave parents "great apprehension." Ex. 597.

Despite this, TikTok took no meaningful steps to mitigate harm. Instead, it continued to privately optimize engagement (and revenue) by deepening the very design patterns its data tied to mental distress. And it concealed its troubling findings, instead publicly promoting "digital-wellbeing."

TikTok's awareness extended beyond addiction and anxiety. Internal analyses and expert feedback directly linked TikTok use to self-harm and suicide. In one 2021 report, TikTok's safety researchers concluded that the risk of users being driven into filter bubbles involving suicide and self-harm material was "HIGH." Ex. 633 at 374:21-376:13; Ex. 517 at 8485. TikTok's internal risk assessments even asked whether the platform was "seeding suicidal and self-harm ideation in young people who had never previously contemplated it." Ex. 599 at slide 42. These internal findings mirrored public health warnings that excessive algorithmic exposure could normalize self-destructive behavior—warnings TikTok ignored. Ex. 1184.

TikTok also received explicit notice from experts about the dangers of the platform promoting eating disorders. [redacted], a public health specialist on adolescent eating disorders, advised TikTok to add warning labels regarding that risk. Ex. 492 at 60:25-70:21; Ex. 659 at 3337; *see also* Ex. 667 at 2. TikTok refused. Instead, it contracted with and paid Eugenia Cooney, an influencer "widely known in the Eating Disorder Community" for severe anorexia, effectively turning her into a spokesperson for disordered behavior. Ex. 492 at 57:11-22, 99:5-24; Ex. 688; *see also* Ex. 492 at 62:21-66:19; Ex. 659 at 3333, 3335-37; Ex. 672B at 343:6-370:22.

TikTok also knew its platform interfered with sleep and learning—two pillars of child development. TikTok's leadership understood that teenagers require 8–10 hours of sleep per night to function properly. Ex. 402A at 134:10-135:8; Ex. 604 at 785. Yet TikTok's internal analyses acknowledged that "social media use at night" is associated with "delayed and inadequate sleep" and that "users are active on TikTok when they should be sleeping." Ex. 632A at 278:10-283:24; Ex. 485 at 0817; Ex. 601 at 9294. Its own data showed that [redacted] of users aged 13–15 and [redacted] of those aged 16–17 were active between midnight and 5 a.m. Ex. 632A at 282:4–285:17. The Trust & Safety team conceded that "the compulsion to react to notifications at night can result in lower-quality sleep," and that this problem "impacts our minor users significantly." Ex. 602 at 1842; Ex. 603 at 8036. Still, TikTok sent notifications to users during the school day and in some cases, up until midnight, which could interfere with sleep. Ex.

540 at 1075; Ex. 603 at 8036 ([redacted] received notifications at "midnight or beyond"). Far from restricting nighttime notifications, or warning young users about their risks, TikTok knowingly disrupted users' rest cycles—the very harm its executives had internally acknowledged.

The same disregard extended to the school environment. Internal discussions admitted that sending notifications during school hours (9 a.m. to 3 p.m.) was "disruptive" and risked "interfering with study at school." Ex. 413B at 202:14-21, 219:25-20:5; Ex. 643 at 8755; Ex. 606 at 2860. Nevertheless, internal analyses—never made public—showed "[redacted]," proving that TikTok's systems encouraged in-class engagement among minors. Ex. 413B at 344:5-10; Ex. 649 at 26-27.

When schools and parent organizations raised concerns, TikTok's response was denial—not reform. Instead of addressing the problem, it instructed the National PTA, one of its paid partners, to send letters claiming TikTok did "not currently see evidence" of school-related challenges—even as TikTok internally tracked over 200,000 videos about the "Devious Licks" challenge disrupting classrooms nationwide. Ex. 444A at 174:23-177:3, 189:12-193:18; Ex. 615 at 4015; Ex. 616 at 5062.

(5) TikTok's failure to exercise reasonable care

TikTok knew its platform was harming children and disrupting schools but chose to do nothing. Internal documents showed students losing sleep, experiencing anxiety, and using TikTok compulsively during class. Rather than address these risks, TikTok concealed the dangers, ignored available safeguards, and prioritized engagement over safety—even as it implemented stronger protections for minors in the Chinese version of TikTok (Douyin) that were never introduced in the U.S. Its intentional and conscious disregard for foreseeable harm has caused widespread classroom distraction, behavioral problems, and drained school resources. As discussed below, this failure to exercise reasonable care—through weak age verification, ineffective parental controls, and other safety lapses—has created a public nuisance that continues to harm and burden the Districts.

(a) Age Verification

TikTok's own policies restrict use of the platform to individuals aged 13 and older, meaning that elementary and most middle school students are supposed to be ineligible. Ex. 400 at 19-20. Nonetheless, it chose not to implement effective age verification measures to prevent these children—including those

using the app at school—from accessing the platform. Ex. 404 at 122:22-123:22; Ex. 403 at 8049 (for age assurance, TikTok relies solely "on an age gate (self-declaration) in all markets").

TikTok's documents confirm that it has long been aware of underage use. Internal estimates showed that at least [redacted] of TikTok's users were under the age of 13, representing more than [redacted] underage accounts in the United States alone. Ex. 418 at 45:6-46:3; Ex. 419 at 2064; Ex. 683A at 110:13-112:14; Ex. 420 at 7491. Other internal reports document an 11-year-old user in 2018, a 7-year-old in 2019, a 10-year-old in 2020, an 11-year-old in 2021, an 11-year-old in 2022, and an 11-year-old in 2023. Ex. 672A at 192:18-193:22, 271:20-272:16, 196:7-197:6, 202:21-204:11, 209:25-211:8, 211:14-213:1; Ex. 650 at 2; Ex. 655 at 2; Ex. 651 at 2; Ex. 652 at 2; Ex. 653 at 2; Ex. 654 at 2.

Although direct messaging was disabled for users under 16, Ex. 402A at 170:5-171:1, TikTok has never verified users' ages—despite acknowledging that age assurance is "the foundation of every Minor Safety and Privacy strategy at TikTok." Ex. 404 at 122:22-123:16, 262:7-264:11; Ex. 401 at 7386; Ex. 403 at 8049 (relying instead on the age gate).

Until February 2019, TikTok had no age gate at all. When TikTok finally introduced one, it merely asked users to self-report their birthdates. Ex. 400 at 18. TikTok employees acknowledged that this measure was ineffective and easily evaded. [redacted], admitted that "all you have to do is look at the age gate and understand that it's easy to put in whatever." [redacted] at 222:8–223:2. Another [redacted], was blunter: "[W]e are idiots to think a kid these days will not cheat the system." [redacted] at 259:23-263:18; Ex. 421 at 2641.

Despite knowing these flaws, TikTok never implemented available tools—such as ID checks, biometric verification, or credit card confirmation—to verify user age at sign-up. Ex. 409 at slide 19; Ex. 404 at 122:22-123:16; Ex. 403 at 8049 (relying solely on the age gate).

TikTok understood that enforcing its own age restrictions would come at a huge financial cost. When [redacted], considered whether TikTok should "remove all U13 [under-13] users," he argued against doing so, citing concerns about losing too many users—and, with them, too much revenue. [redacted] at 186:1-12, 189:18-191:7, 192:21-25; Ex. 478.

The failure to adopt stronger age safeguards is not a matter of speculation or a lack of know-how. Nearly two years before TikTok's U.S. launch, ByteDance introduced the Chinese version of TikTok,

Douyin, in September 2016. Ex. 635 at 94:13-21; Ex. 549 at 2915. The U.S. version of TikTok was built on the same underlying algorithm and core features as the Chinese version (Douyin). Ex. 402B at 480:9-11; Ex. 632B at 450:7-10. The two platforms share nearly identical functionalities—including short-form videos, the "For You" feed, infinite scroll, and push notifications. Ex. 632B at 450:7-17.

A key difference between the Chinese TikTok (Douyin) and the U.S. version lies in age verification and authentication. Chinese TikTok (Douyin) employs an "Age Model" (also referred to as "Age Assurance") to actively verify user ages and trigger interventions that ensure accurate identification of minors for safety features. Ex. 632B at 452:25-453:3; Ex. 496 at 1116 ("Douyin's Stance . . . Important to note they use their age model to deliver interventions"). Chinese TikTok (Douyin) also requires real-name authentication for minor users. Ex. 632B at 467:23-468:6, 572:20-573:12; Ex. 552 at 8104.

The U.S. version, by contrast, does not use the Age Model. Instead, it relies solely on self-reported birthdates through an "Age Gate," a system that depends on user honesty rather than verification. This difference has significant consequences: when minors create accounts using false ages, TikTok's minor safety features are never activated. Ex. 632B at 468:19-469:21.

TikTok fully understood the importance of accurate age determination through the Age Model system. A July 2022 internal document, titled "Minor Safety 2022 H2 Roadmap," acknowledged that "[i]n order to protect minors, we need to know who on our platform is a minor"—a statement with which [redacted], expressly agreed. [redacted] at 471:8-11; Ex. 555 at 4857. In the same document, TikTok admitted that the existing "Age Gate" [in the U.S.] was ineffective because users could easily falsify their birthdates: "relying on age gate turns out to be less effective because there is no way for us to ensure that users are accurately entering their true date of birth." *Id*. at 4862; Ex. 632B at 472:19-473:4. Despite recognizing these flaws, TikTok declined to implement the Age Model used in the Chinese TikTok (Douyin) and continued to rely on the defective Age Gate system in the United States. *Id.*

Rather than strengthening its age gate to prevent underage access, TikTok made it easier to bypass. It introduced "Guest Mode," allowing users to view TikTok videos without creating or signing into an account. Ex. 406 at 3121-3122. TikTok's business team projected that [redacted]

[redacted]." Ex. 407 at 8119. Likewise, TikTok permitted users to avoid the age gate entirely by registering through Facebook or Google accounts. Ex. 697 at 156:21-158:21; Ex. 408 at 6536.

TikTok had the capability to identify and remove underage users through its age inference model but chose not to use it for that purpose in the U.S. For every user, TikTok maintains both a self-reported age and an age inferred by its machine-learning model. Ex. 410 at 23-24. It knew its inferred age data was more accurate and used it when it served its business interests—such as recommending videos or targeting ads based on a user's actual age. Ex. 427B at 617:21-618:24, 621:21-625:8; Ex. 416 at 3077-3078; Ex. 417 at 21-22. Yet in the United States, TikTok deliberately declined to apply that same model to identify and remove suspected underage accounts. Ex. 404 at 124:9-125:17.

TikTok went so far as to mislead regulators about its ability to detect underage users. When asked whether it had any age models capable of identifying or removing such accounts, the company falsely claimed it did not—despite possessing exactly that capability. Ex. 427B at 629:13-632:22; Ex. 480 at 2929A-2930A. Internally, employees acknowledged that if EU or U.S. regulators learned the truth, TikTok might be forced "to delete a large volume of u13 [under-13] users," resulting in an estimated [redacted] loss of its user base. Ex. 427B at 636:18-24; Ex. 480 at 2930A.

TikTok had a clear incentive not to identify children under 13: knowing meant removing them, and removing them meant losing revenue. To preserve its underage user base, the company made decisions to feign ignorance of its users' true ages. In June 2018, as Musical.ly prepared to rebrand as TikTok, [redacted], a member of the rebranding team, proposed surveying users to learn their age. Ex. 433 at 87:8-15, 88:2-89:1; Ex. 568 at 9300. He suggested using education level as a proxy, including whether users were in middle or high school. *Id.* Musical.ly co-founder and [redacted] immediately rejected the idea, noting that including "elementary school" as an indicator would be "definitely violating COPPA," and admitting: "[i]f [the FTC] knows about this survey, they might want us to submit the results, and I am not sure if it's good result." [redacted] at 89:21-90:2, 93:2-15; Ex. 568 at 9301. Another executive, [redacted], warned, "we must avoid any age demographic questions. Doing school level proxy would still be considered age demographics. We take the position in various cases and regulatory matters that we do not collect such information." [redacted] at 95:17-96:22; Ex. 568 at 9301. As a result, [redacted] "made sure none of the age-related questions will appear." *Id.*

TikTok then delayed meaningful action for years rather than remove underage users from its U.S. platform. Its Standard Operating Procedures instructed employees to flag accounts only if a user appeared to be "9 years old" and to ignore parental requests to delete teen accounts. Ex. 697 at 173:2-14; Ex. 481 at 2015. The inaction was not due to technical limitations—it was a choice. When political scrutiny and the threat of a U.S. ban intensified in 2024, TikTok deleted more than 21 million underage accounts in just three months. Ex. 400 at 20. By then, however, the damage to students and school districts was done.

TikTok's willful ignorance extended beyond data collection. TikTok concealed from the public—and from school districts—that children routinely falsified their ages to bypass the platform's restrictions or could evade the age gate entirely by registering through other platforms like Google. Ex. 444A at 262:25-263:4. It also failed to alert schools that they might need to devote resources to curbing student use during instructional hours. Ex. 444B at 365:19-366:6. By hiding the scope of underage use and withholding meaningful safeguards, TikTok deprived schools and parents of the information necessary to protect children—choosing growth and profit over even minimal responsibility for their safety.

(b) Parental Controls

TikTok promoted Family Pairing as one of its signature child safety tools but internally acknowledged that it was deeply flawed. From the outset, it knew design barriers discouraged parents from setting up Family Pairing at all. Ex. 405A at 107:4-10; Ex. 490 at 7039. Even when parents did enroll, TikTok allowed children to unlink their accounts freely—rendering the feature, in its employees' words, "kinda useless." Ex. 632B at 634:10-635:22; Ex. 491 at 6727; Ex. 405A at 48:19-22. TikTok also knew minors were falsely registering as "parents" to bypass restrictions because the system lacked age verification. Ex. 484 at 2445. Parents also received no notice outside the platform itself when their child unlinked from Family Pairing, depriving them of any opportunity to address it. Ex. 405A at 69:18-25.

Internally, TikTok admitted that Family Pairing was rushed and poorly maintained. The feature was pushed through "speedy development" driven by "urgent action in response to regulators," not user safety, and its rollout automatically reset children's privacy settings to "public" when accounts were linked or unlinked. Ex. 405A at 124:25-126:3, 142:9-17, 143:16–144:2; Ex. 493 at 8936, 8939-40; Ex. 494 at 0132. As [redacted], the TikTok leader responsible for Family Pairing, put it, "Family Pairing is where all good product design goes to die[.]" [redacted] at 125:12-18; Ex. 493 at 8939. Despite knowing

these flaws, TikTok never told schools or parents that teens could unlink their accounts or bypass the controls. Ex. 444A at 258:17-23, 259:21-260:4; Ex. 444B at 352:12-21, 353:10-17. Nor did it disclose internal test results showing screen-time management tools were ineffective. Ex. 405A at 199:24-200:22.

In early 2023, TikTok rushed out new "parental control" features under Project M ("M" stands for "Minor"), setting the stage for CEO Shou Chew's polished sales pitch before the United States Congress. Ex. 405B at 318:24-319:21, 333:6-12; Ex. 479 at 4627-28. The Trust & Safety team was instructed to release the tools "by March 23rd since that's when Shou is testifying in front of [C]ongress." *Id*. Internally, Family Pairing was not described as a safety tool or feature but rather as a "PR defensive weapon" (Ex. 489 at 2510) and "a very, very loud way to build awareness around Family Pairing" for regulatory appearances. Ex. 482 at 1893; Ex. 405B at 316:21-317:11.

TikTok's [redacted] remarked that "Family Pairing finally has its moment in the light." [redacted] at 323:2-324:5; Ex. 1187 at 4868. His relief underscored how long the company had postponed safety improvements. The renewed push came only after "executive sponsorship" from [redacted]—reflecting that TikTok acts on safety measures not out of concern for children but when protecting its brand demands it. [redacted] at 324:7-22, 330:7-20; Ex. 1187 at 4869 ("it's coming down from [redacted] and [redacted] so we've got momentum"); Ex. 567 at 7300 ("**It's nothing new, sadly. Just all the things we talked about last year**.") (emphasis added).

ByteDance and TikTok's leadership also made clear that new features could not meaningfully reduce engagement: screen-time tools were permitted to "hurt" metrics by no more than [redacted] for extreme users and minors. Ex. 632A at 278:10-23, 312:4-314:2; Ex. 485 at 0820. When usage dropped, executives questioned "by how much," noting that "[e]ven a few minutes fewer means fewer ads and the impact on revenue at scale is significant." Ex. 405A at 208:20-209:3; Ex. 486 at 7110. TikTok therefore rejected a proposal to adopt a screen-time limit that users could not bypass. Ex. 483 at 5464 (rejecting a hard limit). Internal tests confirmed that these so-called safety tools had no meaningful effect: TikTok found no "negative impact" on minors' average session duration or on key metrics such as daily active users and retention. Ex. 405A at 199:6-7, 199:24-200:5; Ex. 487 at 7228, 7232, 7234; Ex. 488 at 7711. But it withheld these results, admitting "it's not practice to disclose" such data publicly. Ex. 405A at 200:6-22.

In short, TikTok's approach to parental controls prioritized optics—and not losing users—over safety. Although Family Pairing was promoted publicly, including by the CEO in testimony before Congress, as proof of TikTok's dedication to protecting minors, internal documents reveal it was intended primarily as a public relations strategy rather than a meaningful safeguard. The Defendants knew children could easily bypass age and usage restrictions and that its so-called screen-time limits made no meaningful difference. Acknowledging these failures, however, risked reputational and financial harm. Rather than addressing the problem, TikTok chose to protect its image and profits—demonstrating a blatant disregard for the families and schools it claimed to serve.

In promoting Family Pairing as one of its flagship child-safety tools, TikTok publicly proclaimed its noble commitment "to giving parents insight into, and control over," their child's account. Ex. 699 at 3. Yet those assurances are irreconcilable with its internal policies, which deliberately strip parents of the very authority the company claimed to provide. TikTok's own standard operating procedures make this contradiction unmistakable. Employees tasked with communicating with the public are expressly instructed to deny parental requests to delete a child's account if the child is thirteen or older. Ex. 481 at 2015 ("If a parent reaches out to delete their child's account, but the child is age 13–17, we cannot delete the account."); Ex. 698 at 2079; Ex. 697 at 87:8-88:12. Worse still, if a parent persists—renewing the deletion request—employees are directed to ignore the parent entirely. Ex. 698 at 2079 ("If the parent writes back again and gives pushback on the same request of deletion, do not respond.") (emphasis in original); Ex. 481 at 2015; Ex. 697 at 89:17-90:9.

TikTok's disregard for child safety extends beyond families to schools and educators. Reports from teachers or administrators concerning users aged 13–17 are categorically disregarded. Ex. 698 at 2077 (flow chart showing that when a third party—such as a teacher—reports a user aged thirteen or older, no action is taken). Through these policies, TikTok has positioned itself as the ultimate arbiter of what is "best" for teens—overriding the judgment of parents, educators, and others who actually know and care for the child. TikTok's private conduct stands in stark contrast to its public promises of parental empowerment and child safety.

(c) Screen Time Management

TikTok did not implement screen-time management ("STM") tools to curb use, but to counter

growing criticism and reassure regulators the platform was safe. Internal research confirmed that many potential users avoided TikTok because they believed "the platform is addictive." Ex. 632A at 257:21-25; Ex. 495 at 8276. That insight—rather than any genuine concern for user well-being—became the driving force behind the development of TikTok's screen-time features. *Id*.

Internal TikTok documents make clear that the objective of these tools was not behavioral change but managing external perception and protecting the brand and revenue. A high-level product document explicitly stated, "Our goal is not to reduce time spent" on the app, but to influence how users viewed TikTok to increase "DAU [daily active users] and retention." Ex. 485 at 0821. Another described the screen-time controls as "opportunities to improve stakeholders' perception of TikTok as safe for teenagers." Ex. 496 at 1111.

Consistent with that purpose, internal records confirm that TikTok imposed strict limitations to ensure screen-time features would not materially affect engagement or revenue. *See, e.g.*, Ex. 485 at 0820. For all screen-time management projects, [redacted], capped any acceptable reduction in user time at no more than [redacted]. *Id*. ("accept up to [redacted] decrease"). Internal performance analyses further showed that reducing use by merely [redacted] would have exceeded that threshold. [redacted] at 345:20-23. Reflecting those constraints, internal Trust & Safety planning materials show that the team pursued only "moderate" screen-time management initiatives specifically designed to preserve engagement rather than meaningfully curtail it. Ex. 496 at 1112.

When TikTok introduced features allowing users to track their time and set screen-time breaks, it made those tools entirely optional. Ex. 497 at 2841. TikTok's [redacted], acknowledged internally that "anything opt in" receives "[redacted][.]" Ex. 498 at 1898; [redacted] at 181:08-25. Pre-launch testing confirmed as much: only [redacted] of users were expected to use the tools, resulting in less than a [redacted] reduction in time spent on the platform—and no reduction at all for teens—because "minors do not have [the] executive mental function to control their screen time[.]" Ex. 497 at 2841. Nonetheless, TikTok released them anyway—and buried the features behind a "hidden series of menus." Ex. 632B at 593:18–21; Ex. 495 at 8274.

TikTok also introduced "Addiction"—later renamed "Take a Break"—videos to encourage users to log off after extended use. Ex. 499 at 1549. However, those videos were deliberately designed to be

skippable, allowing users to bypass them with "very low friction." *Id*. TikTok's own data showed that over [redacted] of users simply swiped past the prompts, and many left comments mocking the feature's effectiveness. *Id*. at 1549-50.

By late 2022, facing "intensifying criticism at the highest levels of US and EU politicians about addiction-related harm among teens on TikTok," TikTok's Government Relations team urged the company to release a new screen-time feature "quickly in the hope of reducing or redirecting some of this pressure." Ex. 632A at 362:17-363:4, 364:17-23; Ex. 500 at 5806. Although the Trust & Safety team proposed a genuine "hard cap" on teen usage, Government Relations instead advocated for a default limit that "[t]eens can turn off[.]" Ex. 500 at 5805, 5810. The result was a 60-minute default time limit for teens—easily bypassed by entering a four-digit PIN code. Ex. 632B at 537:3-16.

TikTok possessed both the capability and the user feedback necessary to design meaningful screen-time controls but chose not to implement them. In company-run focus groups, half of the teen participants reported that they wanted time-management tools that would prevent them from "simply ignoring [limits] when a limit is reached," such as by "shutting off the app[.]" Ex. 501 at 6136. TikTok, however, only implemented such features in China.

The contrast between TikTok in China (Douyin) and the U.S. platform is stark. TikTok implemented a digital well-being system explicitly focused on addiction prevention on Chinese TikTok (Douyin), aiming to help teenagers "truly and effectively control" their time and "reduce the proportion of users who use Douyin for a long time." Ex. 552 at 8102-8103; Ex. 632B at 459:15-463:8. By contrast, in the U.S., TikTok emphasized user agency rather than addiction mitigation. As one internal document acknowledged, the "[k]ey distinction in our work is Douyin's focus on addiction, while ours is on agency," and TikTok's stated goal in the U.S. was [redacted]. Ex. 552 at 8102; Ex. 485 at 0821.

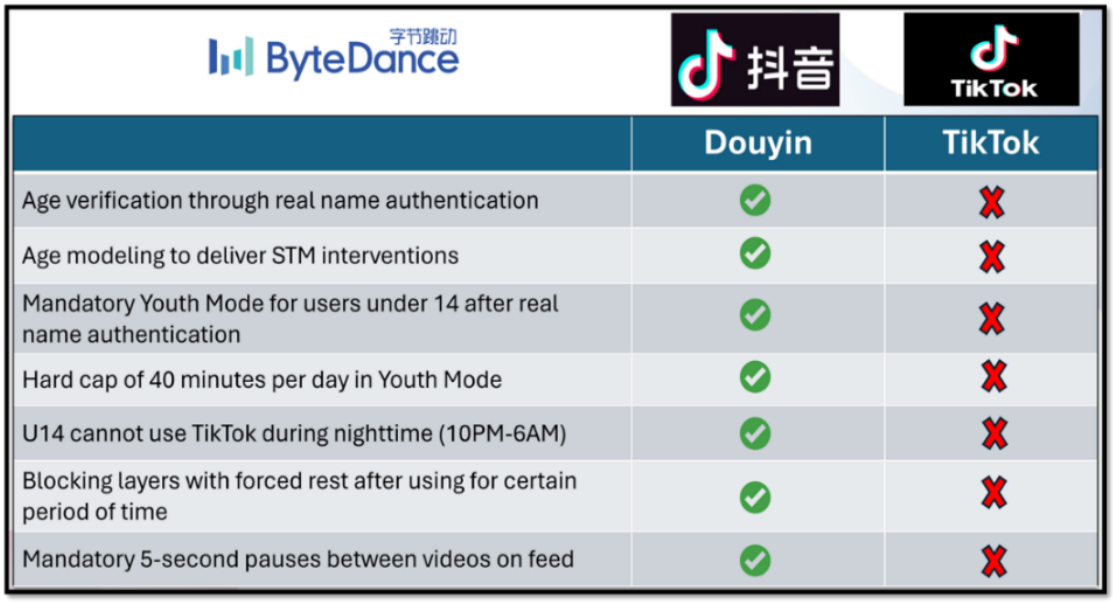
As early as 2021, TikTok implemented strict protections for verified users under 14 on the Chinese TikTok (Douyin), including a mandatory youth mode that imposed a hard cap of 40 minutes per day and prohibited use between 10 p.m. and 6 a.m. Ex. 632B at 473:15-22, 482:10-483:10, 487:11-19. By contrast, TikTok did not introduce its 60-minute daily screen-time limit for minors in the U.S. until 2023—and unlike the Chinese version, in the U.S. that limit was easily bypassed, with no nighttime restrictions of

any kind. Ex. 632B at 411:23-414:11. Internal documents show that TikTok rejected stronger options out of concern for reducing "stay duration." *Id*. Unsurprisingly, its own data later confirmed that excessive use among minors remained widespread: minors averaged [redacted] per day on the platform, and [redacted] were "heavy users" spending more than six hours daily. Ex. 632B at 440:22-441:8; Ex. 544 at 9949.

TikTok was fully aware that minors were using the platform late at night and that excessive nighttime screen use contributed to inadequate sleep and related harms. Ex. 485 at 0816. Internal data from 2022 showed that more than [redacted] of minors were active on the platform between 12 a.m. and 5 a.m. *Id*. at 0817. Despite this knowledge, it never implemented a hard cap or any effective restriction on nighttime use by minors. Ex. 632B at 488:17-489:13.

TikTok also implemented other meaningful safeguards on the Chinese TikTok (Douyin) to limit excessive use among minors—safeguards that were not implemented in the U.S. Beginning in late 2021, TikTok introduced, on the Chinese TikTok (Douyin), blocking layers and forced rest periods that could not be bypassed by users, along with mandatory breaks after extended use. Ex. 632B at 456:1-15 (discussing Douyin's reminder system that "can't be bypassed by the user"), 473:23-476:17 (describing Douyin's "forced rest" periods for users ages 14 to 18 and adults that "can't be scratched away"); Ex. 496 at 1116-1117 (comparing "Douyin's Stance"); Ex. 552 at 8104 (discussing blocking layers); Ex. 554 at 1 (an example of Douyin's reminders). They also added a five-second mandatory pause between videos on the Chinese version for minors—again, a feature users could not skip. Ex. 557; Ex. 580 at 5402. In contrast, TikTok maintained a "user agency" approach in the U.S., relying on self-control rather than mandatory limits, despite knowing that Douyin's stricter controls significantly reduced excessive use. *Id*.

Despite being owned and operated by the same parent company, ByteDance, the Chinese TikTok (Douyin) employs far more restrictive and effective screen-time management measures than the U.S. version, as reflected in Figure 1 below. Technology experts have noted this disparity, describing China's Douyin as a "spinach" version of the app designed to promote healthy use, while the U.S. version is the "opium" version—engineered for addiction and profit. *See* Ex. 506 at 2-3 (quoting Tristan Harris from Oct. 2022).



	Douyin	TikTok
Age verification through real name authentication	✓	✗
Age modeling to deliver STM interventions	✓	✗
Mandatory Youth Mode for users under 14 after real name authentication	✓	✗
Hard cap of 40 minutes per day in Youth Mode	✓	✗
U14 cannot use TikTok during nighttime (10PM-6AM)	✓	✗
Blocking layers with forced rest after using for certain period of time	✓	✗
Mandatory 5-second pauses between videos on feed	✓	✗

Figure 1.

TikTok knew that its screen-time tools in the U.S. were ineffective—both in theory and in practice—yet failed to disclose this information to parents, students, or schools. Ex. 444B at 484:14-490:15. TikTok also possessed both the knowledge and technical capability to build a safer experience for minors in the United States (as was done in China) but chose not to do so. Ex. 402B at 486:3–7, 486:10; Ex. 632B at 581:6-13, 627:3-7 (confirming that nothing prevented ByteDance from launching TikTok with the same protective features implemented on the Chinese TikTok (Douyin)).

In sum, TikTok's approach to screen-time management reflects a deliberate pattern of prioritizing growth and public perception over safety. TikTok knew its tools were ineffective, declined to adopt proven safeguards already implemented on the Chinese TikTok (Douyin), and concealed these deficiencies from parents, schools, and regulators. By choosing to protect its business metrics rather than its youngest users, TikTok failed to exercise reasonable care toward the very minors it publicly claimed to protect.

(d) "The Ugly Side of Beauty Filters"

From the outset, TikTok knew its beauty filters could harm users' mental health and body image, particularly in adolescents—yet chose to proceed anyway. As early as 2018, Musical.ly (later acquired by ByteDance and rebranded as TikTok) faced complaints that its "skinny filter" promoted negative body image, but TikTok deemed those risks "acceptable." Ex. 617 at 0909; Ex. 618. This decision reflected a broader corporate pattern: prioritizing growth and engagement over user well-being. From day one, beauty

filters were embedded into the platform. Ex. 619 at 44; Ex. 632A at 114:10–25.

Despite growing public concern, TikTok ignored repeated warnings that its filters were damaging users' self-esteem, particularly among school-aged users and girls. Media reports throughout 2020 and 2021 raised alarms. Ex. 632A at 184:24-185:6, 196:10-23, 197:12-201:7, 201:14-17, 203:13-205:15, 206:6-207:16, 208:4-210:3; Ex. 620 at 5106. Users themselves pleaded with TikTok to remove or label filters that "trigger people's mental illnesses, body dysmorphia or dysphoria." Ex. 672A at 231:1–3, 232:5-23, 234:9-15, 235:17-238:9, 240:20-242:24, 243:8-21, 245:11-247:7, 248:21-249:4, 249:12-23; Exs. 627-631 (user complaints about the presence of filters).

Internally, TikTok's Trust & Safety team analyzed and wrote a memo about "***The Ugly Side of Beauty Filters***," documenting harms such as "Snapchat dysmorphia," the normalization of cosmetic surgery, and the disproportionate impact on children. Ex. 622 at 4863. A 2022 internal report warned that filters emphasizing thinness, smoothness, or youth could severely damage young users' self-image—specifically identifying "thin," "beautifying," and "age" filters as particularly harmful. Ex. 632A at 191:20-195:6, 195:14-196:4; Ex. 620 at 5102-04.

Despite this knowledge, TikTok did not curtail the use of beauty filters—it expanded it. In 2023, it released a new wave of powerful, AI-driven filters that epitomized the dangers its safety team had warned against. In February, it launched the "Teenage Look" filter, which artificially makes users appear younger, and the "Bold Glamour" filter, which overlays a digital mask of makeup and "perfected" facial features. Ex. 632A at 171:6–20, 172:14–173:16, 174:10–13; Ex. 621. These filters are depicted in Figure 2. *See* Ex. 632A at 179:22-180:5; Ex. 694.

Then, in August of the same year, TikTok went further, introducing the "Body Shrinking" filter—despite long-standing evidence from Musical.ly that such filters distort self-image and promote harmful body ideals, particularly among young, school-aged girls. Ex. 632A at 174:14–18, 181:8-18, 181:15-18; Ex. 695. These filters remain available on TikTok's website today, accessible to users of all ages without restriction or warning. Ex. 632A at 196:1-9; Ex. 623.

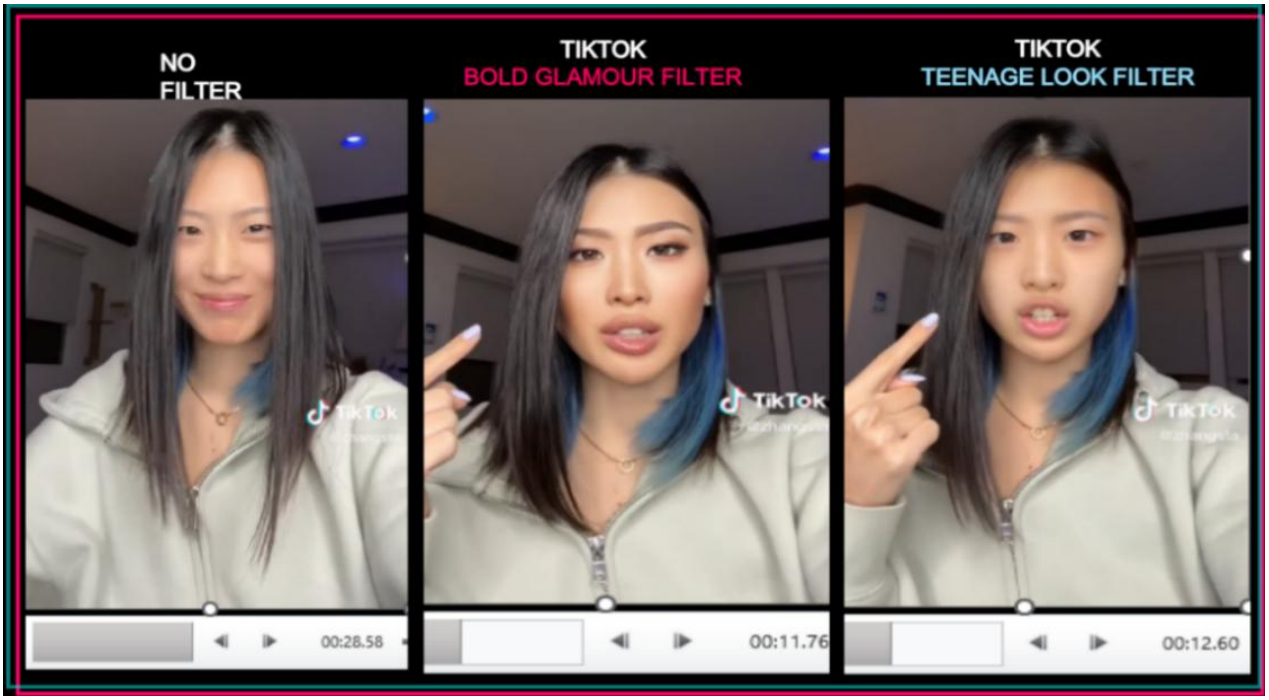


Figure 2.

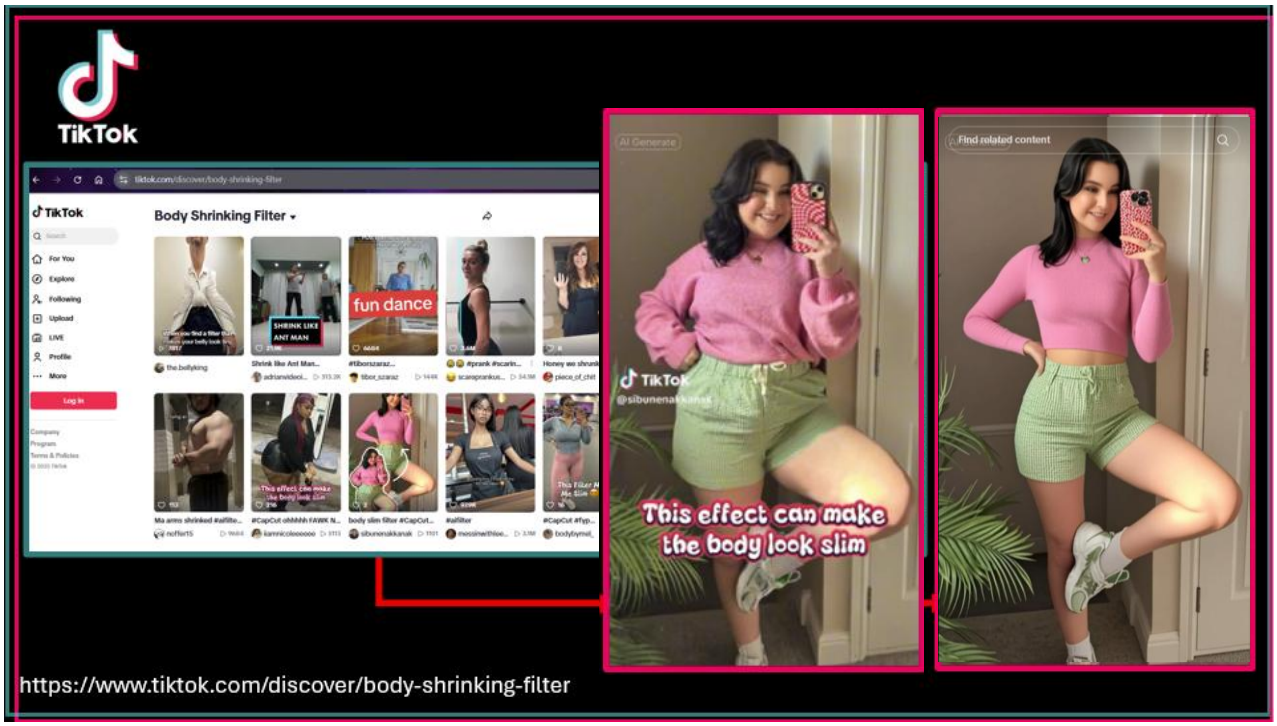


Figure 3.

That same year, TikTok began testing default-on beauty effects—automatically retouching users' faces, including those of minors, without notice or consent. The effects whitened teeth, smoothed skin, and altered facial features such as eyes, lips, and nose. Ex. 633 at 112:5-113:21, 160:13-161:11, 170:8-

171:4; Ex. 624 at 2309-2310, 2315, 2319. These changes were implemented without consulting the Trust & Safety team and in direct contradiction to expert guidance. *Id.;* Ex. 633 at 149:18-151:25.

Inside TikTok, the Trust & Safety team repeatedly warned leadership about the dangers of beauty filters, but its concerns, as always, went unheeded. One member, [redacted], told [redacted] that influencing product teams was "challenging" because safety recommendations "risk[ed] negatively impacting UG [User Growth] metrics." [redacted] at 149:10–14, 152:4–153:20; Ex. 625 at 3295. Another confirmed that new product changes did not routinely undergo safety review. Ex. 633 at 131:24-132:10. By 2024, [redacted] admitted that TikTok lagged industry standards in addressing the harms of beauty filters to young users. [redacted] at 121:4–122:6; Ex. 626 at 7555–56. Yet TikTok never warned parents, educators, or schools about the harms posed by its filters or the safety risks it took in developing them. Ex. 444 at 494:25-498:8; Ex. 418 at 109:21-110:6.

(e) Algorithm

TikTok built its recommendation system for one purpose: to keep users—especially school-aged children—on the app as long as possible. Ex. 531 at 212:18-213:13; Ex. 505 at 0457 (the "main goal" is to maximize "user retention and stay duration"). [redacted] Ex. 427B at 467:1-468:6; Ex. 504 at 8330. [redacted] Ex. 427B at 505:20-506:7; Ex. 503 at slide 21; Ex. 502 at 9725. The formula is simple: more engagement means more ads—and more profit. Ex. 427A at 134:2–13.

In addition to causing compulsive use, the engagement-driven model causes other harms. It traps users in algorithmic "rabbit holes," repeatedly serving similar content based on content-neutral user behaviors ([redacted]). Ex. 507 at 7196; Ex. 427B at 519:23-520:10. TikTok's safety leadership knew these loops caused harm. In early 2020, external experts warned [redacted], that rabbit holes fueled body image issues and even a "contagion effect" around suicide. [redacted] at 217:25-218:15; Ex. 509 at 4763 ("No doubt about it"); *cf.* Ex. 466 (results of "a massive (N = 689,003) experiment on Facebook" in 2014 showing "that emotional states can be transferred to others via emotional contagion"); Ex. 982 ¶¶ 123-125. Although TikTok could identify users caught in these cycles, Ex. 507 at 7199, management

"deprioritized" mitigation efforts and stripped them of resources. Ex. 402A at 245:15-246:10; Ex. 510 at 0555.

After *The Wall Street Journal*'s 2021 exposés, TikTok responded—not to protect users, but to protect revenue. Ex. 633 at 261:12-262:19, 264:14-265:2, 267:25-269:8, 278:14-280:8; Ex. 511 at slide 3. One investigation found that [redacted] of videos served to a new "depression" account related to depression content. Ex. 427A at 416:8-11, 416:25-417:24; Ex. 512 at 0484. Another reported "teen girls sucked into eating disorder/dieting rabbit holes," where TikTok content "triggered" or "accelerated" their conditions. Ex. 531 at 167:23-168:5, 168:17-169:13; Ex. 513 at 0824. TikTok's own data confirmed that [redacted] of eating disorder videos viewed by minors were algorithmically recommended. Ex. 533 at 250:7–20, 252:1-21, 253:14-18; Ex. 515 at 1706.

Advertisers began threatening to pull [redacted], prompting TikTok to act—not for user safety, but to protect revenue. Ex. 402A at 244:8-24; Ex. 510 at 0554 (discussing "[w]hy do we need to solve this problem"). TikTok's [redacted], proposed adding "some diversity" to users' feeds, "but not too much, otherwise users will leave (retention will drop)." [redacted] at 417:25-419:24; Ex. 512 at 0486. TikTok's "filter bubble" fix was capped at a [redacted] reduction in time spent—any more was unacceptable. Ex. 427 at 403:10-25, 427:21–428:13; Ex. 514 at 7749. When eventually implemented in 2022, the solution was only [redacted] effective. Ex. 633 at 281:18-283:16; Ex. 516 at 9678.

TikTok never addressed suicide or self-harm rabbit holes—because *The Wall Street Journal* had not reported on them. Ex. 633 at 270:21-271:5, 276:17-277:5, 278:14-279:4, 374:21-376:13; Ex. 511 at slide 3; Ex. 517 at 8485 ("Ease of entering a filter bubble of [suicide/self-harm] content in U.S." remains "HIGH"). Even after determining with "high confidence" that a user had taken his own life following exposure to suicide content, TikTok did not act. Ex. 633 at 341:18-344:4; Ex. 518 at 5925.

Later, TikTok introduced a "reset" option to refresh feeds, but an internal document acknowledged the flaw: "a [15-year-old] shouldn't have to" rescue themselves from an algorithmic rabbit hole. Ex. 512 at 0485. They also considered a "non-personalized feed" to reduce harm—but implemented it only in Europe, because "ad revenues [would] decrease" if it was offered in the United States. Ex. 519 at 1580.

Despite its knowledge of these risks, TikTok never warned children, parents, or schools that TikTok relentlessly optimized its algorithm for engagement and engagement alone, or that this single-

minded focus on engagement as expressed through users' behavior (time spent, swipes, likes, etc.) could push minors into rabbit holes that put them at risk of eating disorders, depression, and self-harm. Ex. 633 at 386:17-387:18.

Despite its knowledge of these risks, TikTok never warned children, parents, or schools that TikTok relentlessly optimized its algorithm for engagement and engagement alone, or that this single-minded focus on engagement as expressed through users' behavior (time spent, swipes, likes, etc.) could push minors into rabbit holes that put them at risk of eating disorders, depression, and self-harm. Ex. 633 at 386:17-387:18.

(f) Notifications

TikTok's notification system is intentionally designed to drive engagement, not to protect users. Internal documents describe "constant notifications" as a "coercive design tactic" that "detract[s] from user agency[.]" Ex. 541 at 9176. TikTok uses notifications to repeatedly draw users—particularly students—back to the platform. Ex. 411 at 273:2-18 (push notifications appear even when the app is closed); Ex. 411 at 272:18-273:1, 274:10-17 (notifications used "to reengage inactive users"); Ex. 528 at 9604 ("From the platform perspective, push notifications . . . Engage users[;] Drive activeness: directly awake users[;] Lengthen stay duration: better engage users to stay in the app"). The objective was explicit: "encourage users to open the App more and stay longer." Ex. 529 at 9759.

Notifications are not purely automated. Some are manually written and sent by employees of TikTok, and their timing and volume are strategically optimized to [redacted] Ex. 529 at 9765; Ex. 411 at 274:2-9, 282:8-283:2. TikTok actively encourages users to enable notifications, Ex. 565 at 4506, and even when users opt out, many continue receiving them. Ex. 646 at 2 ("I TURNED NOTIFICATIONS ALL OFF! but I'm still getting them!"); Ex. 647 at 2 (user reporting notifications despite disabling them); Ex. 648 at 2 ("I have my notifications completely off . . . and yet I still have notifications").

TikTok is fully aware of the detrimental impact of these tactics on minors. Internal research revealed that children lose sleep due to the barrage of nighttime notifications—as many as 145 in a single evening. Ex. 404 at 83:10-23; Ex. 603 at 8036. Before 2021, TikTok sent notifications during school hours and late into the night, sometimes as late as midnight. Ex. 540 at 1075. The company's own Trust & Safety

team warned that the "compulsion to react to notifications at night" results in poor sleep quality and "impacts our minor users significantly." Ex. 602 at 1842.

TikTok considered limiting notifications to reduce disruptions to minors' sleep and academic performance, including default restrictions during school hours. Ex. 413B at 207:4-9; Ex. 572; Ex. 643 at 8753. Internal analyses warned, however, that such limits would harm "business metrics," causing a [redacted] decline in daily active users. Ex. 643 at 8753; Ex. 572 at 8046. TikTok therefore declined to disable notifications during school hours. Ex. 413B at 274:13-291:25; Ex. 643 at 8753. As one employee admitted, "The only reason" the company did not limit school-hour notifications "is because of the priority." Ex. 413B at 249:1–10; Ex. 573 at 5063. Nor did it implement default settings to restrict use during the school day. Ex. 413B at 198:21-25, 199:1-6.

Despite knowing that these notifications disrupted students' sleep and concentration, TikTok never warned schools, parents, or teachers about their frequency or impact on learning. Ex. 413B at 178:20-179:4, 223:23-224:2, 224:8-13; Ex. 444A at 291:3-292:3.

In short, TikTok weaponized notifications to keep children constantly engaged—during class, late at night, and at the expense of their well-being. Even after internal teams warned that these tactics disrupted sleep and learning, TikTok refused to act, prioritizing engagement metrics over minors' safety. By knowingly maintaining practices that foreseeably harmed students and schools, TikTok once again chose profit over responsibility.

(g) TikTok Now

In 2022, TikTok launched TikTok Now—a feature prompting users to "snap and share front and back camera shots in real time" within a three-minute window after receiving a notification. Ex. 638 at 1183. The company proceeded despite warnings from the Trust & Safety team, which raised "significant concerns from a Minor Safety perspective." *Id.*; Ex. 637; Ex. 640 at 6143 ("[TikTok Now] will diminish our ability to keep teens safe."). The team urged restricting the feature to users 18 and older, but leadership rejected that recommendation, admitting that "[t]he whole point of launching [TikTok Now] is that it is attractive to U18." Ex. 639 at 1944.00006; Ex. 641 at 3556 ("we need to consider business growth in such cases even with lower safety").

Ignoring Trust & Safety, TikTok rushed to release TikTok Now to compete with BeReal, described internally as "the fastest growing product in the history of the internet." Ex. 634 at 4365. Predictably, after launch, minor safety became "the biggest issue risk that surfaced from [TikTok Now]." Ex. 641 at 3556.

Even before launch, TikTok's Trust & Safety team warned that TikTok Now would disrupt classrooms, as it required immediate user engagement following synchronized notifications—often during the school day. Ex. 633 at 392:6-397:3; Ex. 642 at 8648 ("Teens will receive notifications to engage with TikTok Now during school hours."). The team cautioned that "schools and teachers are going to hate" the feature because "kids already have smartphone addiction in class," and adding real-time prompts "with a synchronized push notification and a very real sense of FOMO" would be "a recipe for trouble." Ex. 405B at 359:14-360:5; Ex. 634 at 4364.

Those concerns were quickly realized. TikTok's post-launch data revealed "significant posting" from children ages 13–17 during school hours—triggered by the platform sending TikTok Now notifications in the middle of the school day. Ex. 637 at 1557–58; Ex. 644 at 9718 ("Currently teens receive TikTok Now notifications during the school day"); Ex. 418 at 231:2-17, 233:10-24, 235:9-14, 235:23–236:3 (confirming that responding to TikTok Now notifications in class disrupts instruction). These outcomes mirrored pre-launch warnings that the feature would drive posting "during school hours" and "during sleep hours." Ex. 637 at 1558. But at no point while this feature was being developed or after it was launched did TikTok warn parents, educators, or schools, about the educational disruption TikTok was creating.

(h) TikTok LIVE

TikTok launched its live-streaming feature, TikTok LIVE, around December 2019 as part of a broader effort to expand engagement and monetization among younger users. Ex. 619 at 41-42. TikTok LIVE gave school-aged users—as young as sixteen—the ability to broadcast their lives to a global audience with almost no safeguards in place. *Id.*

Employees quickly saw what that meant. By early 2022, concerns about predation and exploitation on TikTok LIVE had grown so serious that TikTok initiated an internal review—*Project Meramec*—to examine how minors were using the platform and assess the risks it posed. Ex. 418 at 85:22-86:19, 148:6-13. The findings were alarming: children and high school–aged teens were hosting livestreams and

receiving digital "gifts" from adult viewers—gifts that could be converted into real money. Ex. 677 at 65; Ex. 678 at 8273; Ex. 679 at 6385-6386; Ex. 418 at 118:6-120:22, 122:22-123:8, 154:19-161:12, 201:6-205:19. Many of these minors, including those under 13, had accessed TikTok LIVE by signing up through third parties such as Google, Facebook, or Instagram, bypassing TikTok's age gate entirely. *Id.* TikTok's own data confirmed that school-aged users were being exposed to grooming and exploitation—and that the company was profiting from it. *Id.*

[redacted], who led *Project Meramec*, tried to stop it. She warned that TikTok LIVE posed serious dangers to children and teens—ranging from exploitation to addiction—and escalated her concerns directly to senior executives, including [redacted]. [redacted] at 131:23–133:18; Ex. 680 at 1509. [redacted] documented that "senior leaders know about this issue" and that it was "right up front for [redacted] in late February." [redacted] at 138:17-139:11, 149:23-150:2, 150:8-16; Ex. 685 at 8271. Leadership, however, refused to act. As [redacted] later explained, she had "been fighting about fixing this for months with people who didn't think the harms are severe enough." [redacted] at 151:11-19; Ex. 685 at 8271.

TikTok never publicly released the *Project Meramec* data, or warned of the risk of sexual abuse that LIVE created. Ex. 444 at 494:25-498:8. Instead, parents, educators, and schools would have to wait for the press.

In April 2022, *Forbes* published "How TikTok Live Became 'A Strip Club Filled With 15-Year-Olds,'" exposing the same dangers [redacted] had identified. [redacted] at 79:21–82:17; Ex. 682. The article detailed how underage users performed sexual or suggestive acts for adult viewers in exchange for TikTok's virtual gifts. *Id.* at 2. Internally, [redacted] confirmed that "[e]verything this *Forbes* article is addressing are things we already knew" and admitted she had been "on a crusade to fix this problem for months now." [redacted] at 131:23-133:18, 135:4-15, 135:20-136:10; Ex. 680 at 1509. She later acknowledged that the publicity may have been "a good thing" because it "put pressure on leadership to finally take action to protect kids the way [she] was recommending." [redacted] at 140:25-141:4.

[redacted] and her colleague, [redacted], expressed relief that media scrutiny might succeed where internal advocacy had failed. [redacted] wrote, "The only silver lining in press

like this is that they do tend to spark a fire," while [redacted] hoped the article would "put[] a fire behind other potential solutions I've been advocating for." [redacted] at 138:17-139:8, 140:16-142:4; Ex. 680 at 1509. Yet their optimism was tempered by frustration. As another employee put it bluntly: "[P]roduct never takes safety as a priority till things blow up. [N]o amount of complaining helps!" Ex. 418 at 180:24-182:7; Ex. 686 at 1290.

(i) CSAM reporting

TikTok knew it had a serious problem with child sexual abuse material ("CSAM") on its platform. Internal data showed it was "twice as likely to find child sexual exploitation on TikTok than [its] competitors," not because there was more of it, but because TikTok was "facilitating access to CSAM." Ex. 683B at 279:20-280:16, 281:10-282:20; Ex. 687 at 7494. At the same time, TikTok knew it was "under-reporting CSAM" to the National Center for Missing and Exploited Children ("NCMEC"), despite acknowledging internally that "TikTok has a responsibility to report CSAM" to that organization. Ex. 683A at 228:19-229:10; Ex. 683B at 282:23-283:15; Ex. 687 at 7494.00001.

This failure was not accidental—it was the predictable outcome of TikTok's flawed reporting system. Reporting CSAM was so unintuitive that users were effectively discouraged from doing it. In August 2019, employee [redacted] told the company's [redacted], that to report CSAM, a user first had to press the "share" button—an action typically used to send videos to friends, not to report criminal content. [redacted] at 49:18-50:5, 61:2-62:21; Ex. 689 at 2348-49; [redacted] at 231:21-232:10 (confirming "it's not intuitive for a user to press the share button when they come across child pornography"). Even after pressing "share," the reporting option was hidden—"listed on the far right of the sharing options," requiring users to scroll to find it. Ex. 521 at 49:18–50:5, 61:2-62:21; Ex. 689 at 2348-49.

The problem persisted into the following year. Reporting remained "buried under the share button." Ex. 683A at 234:16-235:1. Members of TikTok's Trust & Safety team admitted that the process was "horrible," acknowledging internally that "[i]f we want to be serious about safety and give our words some credence, we need to allow people to easily report problems." Ex. 683A at 230:9-16, 231:2-232:6, 233:3-22; Ex. 690 at 4; Ex. 691 at 1102 ("Clear/intuitive reporting: TT sucks here"). TikTok's leadership recognized the dangers and the failures in its own systems, yet did nothing—ensuring that users remained

unable to easily report serious risks or violations and left parents, educators, and schools in the dark as the problem festered. Ex. 444 at 494:25-498:8.

(j) TikTok's Buried "Safety Center"

[redacted] acknowledged that it was essential for TikTok to provide users, parents, and the public with clear and easily accessible information about how to stay safe on the platform. [redacted] at 227:24-228:11. But when TikTok employees tried to find that information themselves, they couldn't. The Safety Center—the page that was supposed to guide parents and teens through TikTok's safety tools—was so deeply buried that even employees "thought we just didn't have any of this information after ten minutes of searching around." Ex. 683A at 237:24-238:11 (agreeing it is "not reasonable" to expect people to spend that much time searching to "see if TikTok has information on safety"). Internal documents confirmed what everyone already knew: "The additional steps to link to our 'other' help center is awkward and really buries information"; the Safety Center was "not sufficient to meet the needs of our users"; and "[w]e have an obligation to be a lot more on top of this and own our end of the bargain." Ex. 683A at 235:15-237:13, 238:12-239:6; Ex. 684 at 6665-66.

The problem wasn't just neglect—it was a choice. Even TikTok's safety videos, created to educate users about online risks, were "never shown," making them functionally invisible. Ex. 683A at 230:20–231:1. TikTok's supposed safety resources were as hard to find as its willingness to prioritize them.

(k) Barriers to Exit

TikTok deliberately engineered a coercive system to addict young users—and just as deliberately trapped them when they tried to leave. Having built a platform designed to capture attention and drive compulsive use, TikTok ensured that even those seeking to escape its grip would face a maze of obstacles. Every step of the user experience, including account deletion, was crafted to maximize retention and minimize departures, which in turn increased profits.

TikTok's failure to provide effective tools to manage compulsive use has driven many users—especially minors—to try deleting their accounts altogether. Ex. 672A at 159:18-20, 160:4-10, 162:2-13, 167:24-168:5, 169:2-7, 175:9-17; Ex. 583 at 2 (user reporting, "it got me addicted and now I have to delete it"); Ex. 586 at 2 ("I've come to the decision to delete my account"); Ex. 587 at 2 ("Trying to get over my TikTok addiction. Please delete."). But TikTok intentionally made that process difficult. Deleting an

account requires navigating through ten separate steps—a system deliberately engineered to discourage users from leaving. Ex. 669 at 17 (images depicting the ten steps needed to delete an account).

In 2021, TikTok received roughly [redacted] account deletion requests per day. Ex. 670 at 9388. Rather than viewing this as evidence of user harm, it treated it as a retention problem and added new barriers "to make sure only users who really want to delete [an] account submit the deletion requests." *Id.* TikTok implemented two deterrents: a "pre-check" reminder prompting users to complete tasks before deletion, and a "questionnaire" asking why they wished to delete their accounts. *Id.* These tactics worked as intended—by December 2021, [redacted] of users quit the deletion process midway, most abandoning it at the questionnaire ([redacted]) or data download screen ([redacted]). Ex. 671 at 9722.

From those questionnaires, TikTok learned that many users sought only temporary deletion—often citing "temporal leave" ([redacted]) or "other reason" ([redacted])—and immediately began exploring ways to "save" those users by offering temporary deletion options and re-engagement strategies. *Id.* Even users who completed all ten steps could not easily leave. TikTok automatically placed deleted accounts into a 30-day reactivation period, during which logging in even once canceled the deletion process. Ex. 672A at 311:20–312:2, 313:10–22; Ex. 645; Ex. 669 at 13. TikTok knew users struggling with addiction were especially vulnerable during that time. Ex. 672A at 310:10–19, 311:2–15, 314:8–315:15, 322:6-19, 322:23–325:1, 326:25-327:16, 327:20–328:15; Ex. 673 at 2 ("I have already deleted the account several times."); Ex. 674 at 2 ("Don't give me 30 days because I'll always come back to it."); Ex. 675 at 2 ("Please delete my account immediately. I am unable to wait the 30 days because I am addicted to this app."). Yet even when users explicitly requested immediate deletion, TikTok refused to shorten or waive the reactivation period. Ex. 672A at 314:3-5.

(6) TikTok's misrepresentations and manipulations

TikTok's leadership actively buried the findings of the company concerning user harm. TikTok maintained [redacted], while publicly issuing statements claiming that its platform promoted "creativity, self-expression, and joy." Internal studies were never released, and public-facing materials like the "TikTok Guide for Parents" made no mention of the known and documented mental health, sleep, or educational harms.

TikTok's consistent response to evidence of harm was to manage perception, not risk. It

commissioned studies to identify problems, then silenced or reframed the results to protect engagement metrics and corporate image. In doing so, it perpetuated the very harms its researchers—both internal and external—warned of, proving that these consequences were not only foreseeable but foreseen.

(7) TikTok failed to warn parents, educators, and schools about known mental health risks

As detailed above, TikTok knew that its platform placed children and schools at grave and foreseeable risk. It knew that school-aged users were being harmed by compulsive use, sleep loss, body-image distortion, eating disorders, depression, anxiety, and exposure to self-harm content. Yet it chose silence—failing to warn the very parents, educators, and institutions responsible for protecting those children.

TikTok failed to warn students, parents, educators, school administrators, and school districts that use of TikTok may be harmful to the mental health of school-aged children. Ex. 413B at 173:3-14. Nor did TikTok warn these groups that use of the platform increases the risk of serious harms, including addiction, compulsive or excessive use, depression, anxiety, suicidality, self-harm, eating disorders, body dysmorphia, and sleep disorders. Ex. 413B at 172:7-11, 172:16-173:1, 173:14-21, 173:23-174:4, 174:6–12, 175:8-15, 175:17-25, 176:2-7; Ex. 444A at 293:18-296:1.

Similarly, TikTok failed to warn teachers, school administrators, and school districts that students' use of TikTok foreseeably harms schools. Ex. 413B at 177:8-13; Ex. 444A at 289:18-290:1. TikTok did not disclose that schools would likely need to divert staff time and resources to: (1) confiscate cell phones used for TikTok during school hours; (2) respond to threats made against schools and students through the TikTok platform; and (3) implement enhanced technology measures to restrict student access during school hours. Ex. 444A at 296:24-297:8, 297:10-17, 298:3-11. Nor did it warn that student that TikTok use during class time disrupts instruction and distracts both students and teachers. Ex. 413B at 177:15-20, 178:4-13; Ex. 444A at 290:2-9.

Moreover, TikTok failed to inform schools that students can post on TikTok and receive notifications during school hours—and that such notifications interfere with study and concentration, an issue TikTok's own teams had already recognized. *See supra* § (f) (Notifications); Ex. 413B at 178:20-179:4, 223:23-224:2, 224:8-13; Ex. 444A at 291:3-292:3. Finally, TikTok failed to warn parents,

educators, and schools that, as a foreseeable result of student TikTok use, schools would likely need to: (1) hire additional mental health staff; (2) train staff about the harmful effects of social media; (3) develop educational materials addressing social media addiction; (4) divert staff time to confiscate cell phones during school hours; (5) repair property damage linked to TikTok-related incidents; and (6) increase technology resources to restrict access to TikTok. Ex. 444A at 292:5-293:16, 296:3-9, 296:25-297:8, 298:3-11.

Despite knowing the harms its platform inflicts on students and schools, TikTok continued to market itself as an educational tool, boasting that "millions of people learn something new on TikTok." Ex. 696. Its ads—such as the one featuring "Mrs. Kelly" depicted to the right (*Figure 4*)—underscore the deception, celebrating supposed educational value while concealing the platform's harm to children, educators, and schools.

c) Snap's Conduct

(1) Snapchat's business model

Snapchat is a social media platform that catapulted to popularity with teens shortly after its launch in 2011. Ex. 800 at 7142 (Snapchat users are "primarily a younger demographic"). Snap generates revenue primarily by advertising to its users. Ex. 829 at 51:14-18 (advertising was Snap's principal method of generating revenue since 2014). Snapchat advertisements appear across the app, in between posted Stories, on the public Discover and Spotlight video feeds, within the camera, when browsing augmented reality lenses, and through the Snap Map. *See* Ex. 921 at 1031 (slide depicting where Story Ads appear); Ex. 807 at 39:17-40:11, 81:20-82:14 (describing appearance of ads in Discover feed and on Spotlight); Ex. 1150 at 7128 (ads appear in Stories, camera, Discover, and in commercials); Ex. 829 at 54:3-14 (ads appear in Snap Map). To increase revenue, Snap aims to increase daily active users, increase engagement, and to increase the number of active advertisers. Ex. 921 at 1043.

Snap ads are specifically designed to increase daily user time. *See* 1151 at 0291 (design plan explaining that "ads will be more closely linked together, making a better user experience for all and

keeping Snapchatters on the platform for a longer period of time"). Snap also created profiles of its users so advertisers could target by age, gender, location, language, lifestyle, shopping preferences, and numerous other characteristics. *See* Ex. 807 at 153:16-162:24. By 2018, Snap offered advertisers self-service interfaces so they could seamlessly create ads using these characteristics without interfacing with a Snap representative. *See id.* at 24:7-16.

Teens are core to Snapchat's business model. *See* Ex. 827 at 112:13-24 (head of product research: "our core audience is teenagers."); Ex. 807 at 127:11-16 (former product team lead: "[Snapchat is] the number one social media app for teenagers in the United States."). Snapchat dominates in terms of teen market penetration—by 2014, Snap boasted that it "absolutely own[s] the youth demo in the US," with roughly 84% of the market among users aged 10 to 24. Ex. 825 at 4207. Ahead of Snap's IPO three years later, investors observed that middle school children under the age of thirteen were "rabid . . . user[s] of Snapchat." Ex. 826 at 7064; *see* Ex. 866 at 7658-59 (email sent from Snap's head of communications to CEO reporting that 40 percent of children under 13 use Snapchat). By 2023, this number had grown, with Snapchat reaching "90% of the 13- to 24-year-old population[.]," Ex. 822 at 4550. That same year, CEO Evan Spiegel estimated that about 90 million users were teens. Ex. 803 at 140:21-141:7. (As discussed below, this number was likely far higher given that many teens report their age inaccurately.)

Further, users aged 13-17 represent nearly 50 percent of the time spent on Snapchat. Ex. 833. Internal documents characterized Snap as a "daily habit" for almost 50 percent of "extremely engaged" 13-to-17-year-olds in the U.S. Ex. 823 at 1034 ("Young users are extremely engaged and send Snaps daily"); *see* Ex. 831 at 6185 ("this generation can't live without our app"). And, as Snap's vice president of Product testified, users spending more time on the app generates more revenue. *See, e.g.*, Ex. 829 at 63:17-64:17; Ex. 828 at 7675 (Snap "revenue formula" turns on "time spent" per monthly active user).

Snap has powerful incentives to capture users as young as possible. Internal analyses showed that the 13-to-17 age cohort generated the highest average revenue per user (ARPU) during the first 12 months following registration. Ex. 828 at 7745-746; Ex. 829 at 75:24-77:6. The economic advantage of early acquisition is clear: a senior product designer testified that acquiring a user at age 13 would "produce revenue for Snap at age 13, 14, 15, 16, 17, and 18, and then forward," whereas users joining later in life generated much less revenue. Ex. 829 at 80:5-81:2 (agreeing that "it's to Snap's benefit to onboard users

as early as possible to maximize the amount of time that a user is in this 13 to 17 demographic where they are . . . the highest revenue-producing demographic on Snapchat"); *see* Ex. 830A at 215:18-21 (Senior Director of Product Management: teenage users who "start using Snapchat when they're 16 or 17, a really high proportion of them continue using that service"). Internal presentations explicitly emphasized the importance of targeting the youngest allowable users, labeling the cohort "very valuable for advertisers" and emphasizing that the "Snapchat Generation has $2 trillion in purchasing power, making up over 40 percent of global consumers." Ex. 831 at 6185; *id.* at 6182 ("continuing to win with new 13 year olds is the most critical aspect to onboarding new users"); Ex. 829 at 104:11-105:5; Ex. 991 ¶¶ 154-156. Snap also claimed that, through teens, advertisers could reach their parents, because the "Ads [teens] saw on Snapchat influence what they asked for and what their parents purchased." Ex. 902 at 8460.

Snapchat has seen tremendous success. Revenue increased by an average of 21.2% per year from 2020 through 2024, hitting $5.4 billion in 2024. *See* Ex. 836, Ex. 837, Ex. 838, Ex. 839, Ex. 840, Ex. 841 (10-K and 10-Q Reports). Importantly, a significant percentage of Snap's profitability is attributable to teenage and adolescent users; from 2020-2023, Snap earned over $1 billion in revenue each year from the under-18 age cohort. Ex. 991 ¶¶ 164-65.

(2) Snapchat's targeting of school-aged children

Not surprisingly given its business model, Snapchat is designed to appeal to teens and adolescents. Ex. 817 at 3471 (internal research explaining that "childish perception of Snapchat derives from the stereotype that that [sic] most Snapchat users are young people"). The platform appears in bright colors, primarily yellows. Its logo is a cartoonish ghost. Ex. 818 at 9163 (adult users observe that "the overall aesthetic of the app feels cartoonish"). Further, users are represented on the platform via "Bitmoji"—avatars with childlike proportions that allow each user to present themselves as a cartoon version of anyone they choose. Ex. 819 at 5958 ("Bitmoji Stories have become the Sunday cartoons"). Figure 5 is a representative depiction of Snapchat's bitmoji-driven aesthetic, from its internal documents (Ex. 815 at 0094). Snap uses Bitmojis to infer user gender for "content and ad ranking, suggestions, targeting, [and] optimization." Ex. 820 at 5893 (describing plan to use Bitmoji data to improve gender inference for advertising and noting "[g]ender is one of our strongest user profile attributes"). During the school year,

Snap's goal is to "[k]eep Bitmoji top-of-mind among teen-skewing lifestyle media and introduce them to the many ways they can use Bitmoji to personalize their back-to-school style." Ex. 821 at 2488.

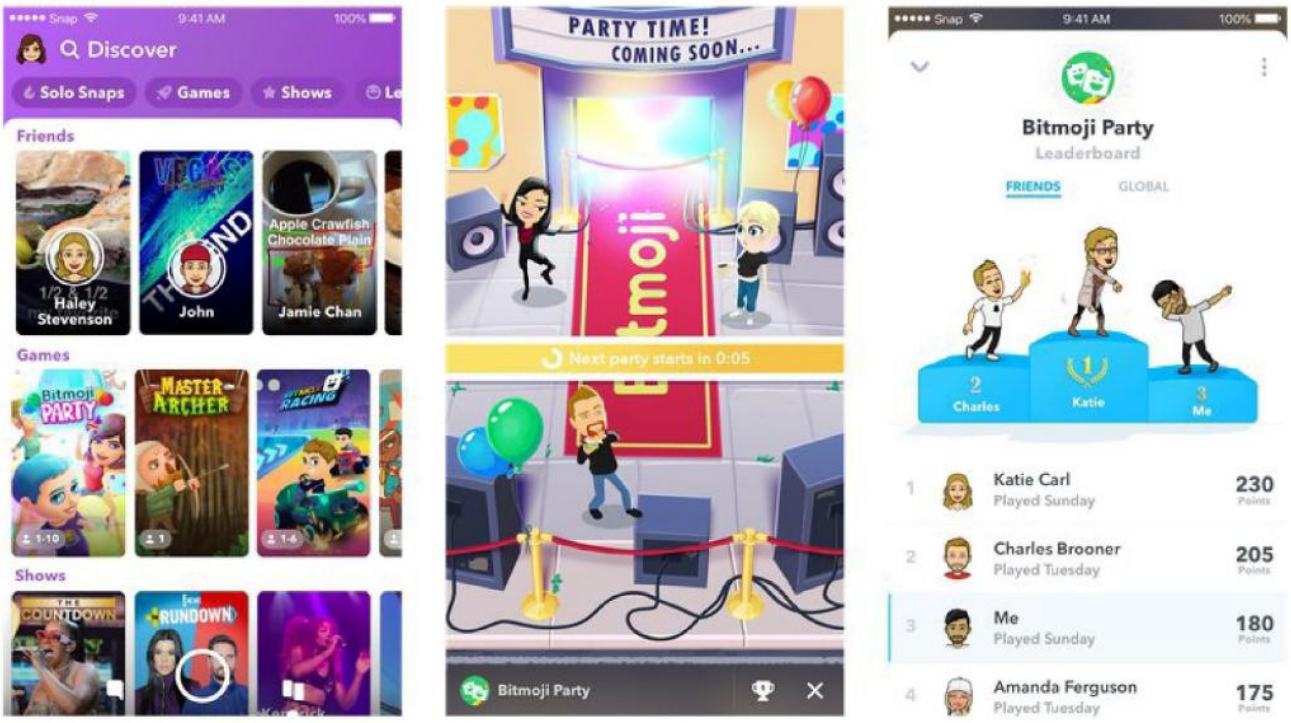


Figure 5 (from Ex. 815 at 0094).

Snapchat also makes use of numerous other features that appeal primarily to teenage sensibilities, including ephemeral messaging, gamified features like Streaks, and augmented reality lenses (some of which are playful and cartoonish). Upon launch, Snapchat's most distinctive characteristic was that video or photo messages, or "Snaps," were ephemeral, and vanished shortly after they were received. Ex. 801 at 661:5-663:10; Ex. 802 at 0241; *see also* Ex. 803 at 89:18-90:22; Ex. 804 at 5803 (Snapchat was created as an app whose primary purpose was to allow users to send each other photos that would automatically disappear). Snap touted its disappearing messages as a way for "young people" to "combat" the pressure to "curat[e] a perfect image of themselves on social media." Ex. 816 at 2006. But the dark side of this feature was made clear almost from its inception—Snapchat "quickly gained a reputation as a sexting platform." Ex. 805 at 4149.

While other platforms copied Snapchat's signature features like ephemeral messaging and Bitmoji, Snapchat itself evolved by adopting key characteristics common to other major social media platforms. *See* Ex. 834 at 96:25-97:17 (referencing numerous features common to other social media platforms). As

with Instagram and TikTok, users are encouraged to enhance their appearance in video or photos by applying beautification filters to smooth or whiten their skin, remove blemishes, and make them appear slimmer. *See* Ex. 835 at 0990 (describing lenses to include: "Skin smoothing/whitening," "Acne/Blemish removal," "Dark circling / Eye enlarger / Eye brightening," and "Slimming / Stretching"). At the same time, many of Snapchat's lenses, like the platform's overall aesthetic, are overtly cartoonish, for example, allowing "young people to put on bunny ears or [] vomit a rainbow." Ex. 801 at 466:10-467:1; Ex. 865 at 6224 ("Snapchat lenses are believed to be more childish than TikTok lenses overall"). Users are also presented with an algorithmic video feed that mimics TikTok's endless scroll. *See* Ex. 808 at 3050 ("We already have an endless scroll design in Discover and I think we wish it was more ~~addictive~~ compelling.") (strikethrough in original); Ex. 807 at 311:21-312:1 (acknowledging similarities between TikTok and Snapchat's Spotlight); *see also* Ex. 834 at 96:24-97:17. And the videos they post can receive "likes" or "favorites" that signify the social approval of their peers. *See* Ex. 807 at 297:20-22.

Snap has innovated on the idea of social communication by transforming it into a gamified experience—encouraging users to send daily Snaps to friends, acquaintances, and strangers in their friend list simply to maintain a "Snap Streak" with those users, *see* Ex. 801 at 193:22-194:3; share with other users where they are—and what they are missing—at all times via a map feature, *see* Ex. 809 at 8084 ("problem statement deck" identifying "FOMO – seeing a party happening on Snap Map" as a "negative social interaction"); and broadcast content to friends by posting video Stories, *see* Ex. 807 at 37:17-40:24, which, like Snaps, are ephemeral and thus prompt users to constantly check the platform before Stories disappear. *See* Ex. 1152 at 2064 ("The ephemeral nature of Stories requires users to develop a daily habit of checking them"). Snap provides each user with a "Snap Score" that awards more points for greater engagement with the app. Ex. 810 at 3228 (describing Snap Score as "feeding into the gamification mindset among Snapchatters"); Ex. 811 at 4720 (Score functions to "reward and incentivize users" for virtually every platform interaction—sending Snaps, adding friends, posting Stories, or maintaining Streaks). Snap Trophies rewarded users for achieving Snap Score milestones, engaging in large volumes of chats, utilizing certain filters, or viewing prescribed amounts of video content on Snapchat's endless video feeds, Discover and Spotlight. Ex. 814 at 4019-20, 4053-60. And users are incentivized to add an ever-expanding cohort of "friends," most of whom they do not know in real life. *See* Ex. 806 at 6130

(teens use group "Shout Outs" to increase "the visibility of their profile" so that "loads of people" will "add them" as friends). Snap added this last feature to encourage users to further increase their Snap Score, enmeshing each user in deeper engagement with the platform.

Snap understands that Snapchat's interconnected reward systems create addictive usage patterns disproportionately affecting adolescents. Ex. 810 at 3227 (acknowledging that Snapchat "is social gamification, which tends to be popular amongst the younger demographics."). Snap expressly recognized the psychological impact of the Snap Score, for example, acknowledging that "[e]ducating users" about their Snap Score creates an "addictive" effect that "encourage[s] them to produce and consume more snaps," transforming them into what Snap calls "high-quality users." Ex. 812 at 8652. Snap also understood that the gamified Score presented safety risks, incentivizing young people to "keep their privacy settings open, accepting strangers' [friend requests], and participate in shout outs and streaks in order to gain more contacts and views on Snapchat and eventually increase their score." Ex. 983 ¶ 93.

Given its numerous design choices to target teenagers, Snapchat struggles to attract and retain users who are in their mid-twenties and older. *See, e.g.*, Ex. 823 at 1034 ("In US, Snap sending is an everyday . . . behavior for young users – but not so for older users"); Ex. 830A at 190:1-191:3 ("Gen X and older users generally have never used Snapchat"). Thus, in order to maintain growth, Snap must constantly drive users to Snapchat while they are young to offset the loss of users who age out of the platform's teenage-focused ecosystem. Senior Snap executives have acknowledged the importance of continually recruiting new teenage users— with Spiegel explicitly asking, "Should we acquire new 13 year old users?" and CFO Derek Anderson emphasizing that continuing to "win with new 13yr olds" is "super key" and "the most critical aspect to onboarding new users." Ex. 831 at 6182-83. Conversely, when confronted with data showing "a decline in the US 13-to-17 trend line," Boyle wrote that he felt "absolutely sick to my stomach," while Spiegel and other executives expressed "great concern" at reports of "a dramatic drop in 13 to 17" year-old users. Ex. 842; Ex. 843 at 7523; Ex. 834 at 28:4-21.

(3) Snapchat's targeting of schools

The financial imperative to relentlessly recruit young users led Snap to deliberately target and infiltrate schools. Snap understood that schools provided a crucial avenue for growth. High school students

in particular were characterized by Snap as its "core target audience" and "instrumental" to growth. Ex. 855 at 5097, 99; *see also* Ex. 856 at 1476. During back-to-school periods, Snap observed large spikes in new friendship, finding that users add "47% more friends from Group Chats during back to school months." Ex. 857 at 0369. This surge early in the school year provided "a tailwind for engagement for the remainder of the year," and led to sustained use throughout the school year. *id.* at 0359-60.

Accordingly, Snap leaned especially hard into the back-to-school period, consciously seeking to "make Snapchat synonymous with back-to-school" and, specifically, the "purchasing behaviors" associated with that time of year. Ex. 867 at 97:15-98:2; *see also* Ex. 830A at 246:5-247:6 (acknowledging that Snap considers back-to-school season to be a "really important time of year" and that it is a "commonly understood period in time where there's additional marketing"). Snap created "back to school" messaging to convey to teenagers and adolescents that Snapchat "is the app for students." Ex. 860 at 0064. Snap claimed to be "part of the family during back-to-school." Ex. 902 at 8443. Snap also strategized on how to "[g]et High Schoolers attention" by leveraging its features, such as filters and lenses, to create a "high intent audience," as "over 90% of high schoolers are in the lens carousel on a monthly basis." *Id.* at 8472, 74, 79. The lens carousel included school-related offerings such as a "School Vibes" and "School Style" filter. Ex. 904; Ex. 905.

Meanwhile, Snap told advertisers that teenagers used Snapchat for every action leading up to their back-to-school purchases: 21% looked for inspiration in Stories, 33% used the app to get advice from friends, nearly 75% were interested in "trying" new products using Snapchat's augmented reality features before buying, and 27% took Snaps using in-store filters. Ex. 902 at 8468. Snap created numerous additional avenues for advertisers to reach teenagers at school, including full-screen Snap Ads and commercials on the Discover tab, each with the goal of creating "instant impact" and to "start building a qualified audience" of teenagers at school. Ex. 903 at 0337. And, as described in more detail below, by 2016, Snap obtained precise high school location data to enable advertisers to geotarget students on campus, boasting to advertisers that Snapchat "rules the halls of high school." Ex. 807 at 134:24-135:7, 138:19-140: 24; Ex. 903 at 0334; *see also* Ex. 903 at 0032 ("Snapchat reaches 2x as many 13-17 year-olds in the US than Facebook, Instagram, and Messenger – combined").

Snap's intentional targeting of schools is woven into its history. Shortly after Snapchat's 2011 launch, Evan Spiegel wrote in an official company blog post that he was "thrilled to hear that most of [Snapchat's users] were high school students who were using Snapchat as a new way to pass notes in class—behind-the-back photos of teachers and funny faces were sent back and forth throughout the day." Ex. 824 at 2; Ex. 803 at 111:14-113:14. Spiegel added that the app's usage skyrocketed during school hours and otherwise waned. Ex. 824; Ex. 803 at 111:14-113:14.

In 2017, Snap's Chief Strategy Officer provided Spiegel with an investor analysis titled, "Snap: A Middle School Crush," which described "the middle school set [as] a rabid – almost exclusively – user of Snapchat." Ex. 826 at 7064. Internal company documents from 2018 showed that 64% of Snapchat users aged 13-21 reported using Snapchat during school. Ex. 845 at 3471. This was consistent with outside research performed on Snap's behalf, which showed that Snapchat accounted for 39-40% of the total battery usage for high school students, while it accounted for 31% for middle school students. Ex. 846 at 5372-73; *see also* Ex. 1000 ¶ 39 (explaining that 45% of phone use at school "was for Snapchat alone"). Snap also surveyed teenage users and knew that teens "use Snapchat to communicate with friends" at school "unless the teacher tells them not to." Ex. 849 at 7235; Ex. 827 at 470:10-472:2. Snap personnel referred to students' Snapchat use during school as "under-the-desk" phone time, and Snap witnesses confirmed they were aware that students used Snapchat on their phones during class and took pictures of events at school. Ex. 851 at 416:10-417:24; Ex. 852 at 6439; Ex. 853 at 367:14-368:1. In short, Snap has long known that Snapchat is frequently used in classrooms during school hours across the United States, and it encouraged that use. Ex. 854 at 1969.

To further boost student engagement, Snap created a feature called School Communities to connect students at the same school by class year. Ex. 857 at 0372. Each School Community was a closed space where students could share school-related Stories and send messages to one another. Ex. 861 at 124:19-125:20. The point of the School Community feature was to "grow[] the user base" by targeting high school students and providing them a place to congregate. Ex. 864 at 8680. At a 2023 company town hall, Snap personnel boasted that its School Community feature had "15M students across 24K US high schools!" Ex. 857 at 0371; 861 at 125:1-20. Because of its growth potential, School Community caught the eye of Snap's competitors, with Meta personnel internally remarking, "Our competitors in social connection are

leveraging school integrations as a way to keep in touch with their community." Ex. 258 at 7145; *see also* Ex. 269 at 5745 (Instagram presentation focusing on School Communities and observing, "Our competitors are exploring school communities").

Following rollout of this growth-centric feature, Snap "learned that some of the largest public school districts ban external emails, preventing students from receiving the identity verification emails" from Snapchat, which were necessary for students to join their School Community group. Ex. 864 at 8679. Rather than respect these external email bans, Snap created alternative verification methods to circumvent the email restriction. *Id.* Likewise, Snap developed technology to circumvent other proactive efforts by schools to ban or limit Snapchat use in school. During COVID, when students were participating in web-based learning, senior Snap personnel noted "under the desk" phone time was decreasing because children weren't in school, and recommended that the company turn its attention to developing a web-based product. Ex. 851 at 417:3-419:11. At Spiegel's direction, Snap developed Snapchat for Web, allowing students to access the platform from a laptop or tablet to avoid discipline for using smartphones during class. Ex. 934 at 6632 (explaining that Spiegel directed the development of a "fun and engaging laptop experience" so that kids could use Snapchat "in class"). In addition, Snap prepared for schools to eventually ban cellphone use altogether, designing a wearable device on which students could use Snapchat during class should they be prohibited from using their phones. Ex. 952 at 4134-35. These deliberate workarounds show that Snap not only knew its platform disrupted schools but actively engineered ways to ensure that disruption continued. It is especially ironic that Defendants now cite COVID as an alternative cause of school disruption when Snap's own records show it treated the pandemic as an opportunity to deepen student engagement and maintain classroom use.

(4) Snap's knowledge and willful blindness of harms

Although Snap's relentless pursuit of teen engagement has been good for the company's growth, Snap knows that use of Snapchat can be harmful for minors. And while the relentless penetration of Snapchat into schools has likewise helped Snap's bottom line, Snap understands that use of Snapchat at school is disruptive, interferes with the educational environment, and harms schools. Snap has long known enough that it should have warned users and their caretakers, including parents and schools. Instead of doing so, Snap went out of its way to maintain a stance of plausible deniability.

Snap's co-founders, Spiegel and Bobby Murphy, have been aware since as early as 2013 that users experience what they termed "Snapchat addiction." Ex. 922 at 4154 ("frankly the ones that have the Snapchat addiction have no room for anything else. Snaps dominates their life."). Over the years, young users regularly reported to Snap that they were distressed from their compulsive use of Snapchat, with some reporting they felt addicted and admitting to opening the app several hundred times per day. *E.g.*, Ex. 928. In 2018, employees casually discussed articles describing Snapchat's "elements of addictive design [that] encourage high user engagement." Ex. 923. These articles were consistent not only with user reports but with Snap's internal findings that Snapchat had become a "daily habit" for nearly 50% of "extremely engaged" 13- to 17-year-old daily active users in the United States. Ex. 924 at 1034.

Snap understands that addictive and habitual use is driven in part by specific Snapchat design choices. For example, because the Snapchat platform is built around ephemerality, it is designed to "drive a phenomenon commonly called FOMO—Fear of Missing Out. Knowing that a funny video or a party photo will vanish by tomorrow compels teens to open the app frequently, so they don't 'miss' what their friends are doing. It creates a 24/7 urgency to be online." Ex. 989 ¶ 84; Ex. 927 at 2045 ("The ephemeral nature of Stories requires users to develop a daily habit of checking them.").

Meanwhile, Snapstreaks and the Snap Score gamify the Snapchat user experience in a way Snap understood to be addictive and anxiety-provoking. *See* Ex. 812 at 8652 ("Educating users what score is and how to increase score may make users care about (if not become addicted to) their score and thus encourage them to produce and consume more snaps, and ultimately become high-quality users"). The Snapstreak feature, launched in 2015, requires users to exchange Snaps daily with specific friends to maintain numerical counters displayed with fire emojis; if a user fails to send a Snap on a given day, the Streak is lost. Snap recognized the feature's compulsive nature, characterizing this mechanism as "Daily Habit Training." Ex. 925 at 6161. Further, it acknowledged that the Snapstreak feature "tapped into mass psychosis" where "people *must* keep the streaks going." Ex. 935 at 9344; *see also* Ex. 987 ¶ 143 (describing Streaks as Snap's "'playing by appointment,' compelling users—especially teens—to return to the platform daily in order to avoid losing [the Streak]"). Internally, Snap's product team expressed the "stance on streaks…that we don't love them (it was an accidentally addictive, somewhat unhealthy feature that gamifies friendship in a weird way)." Ex. 926 at 9271. "[B]ut they're too delicate to touch right

now"—given that over fifty million users had Streaks. *Id.* In 2018, Spiegel himself asked senior product managers for a plan "to kill streaks" because "it's the right thing to do." Ex. 971 at 6015. But Streaks was never killed, because of its impact on teen engagement. Ex. 1153 at 1495-96 (referring to streaks as "compulsive behavior" that is "a main driver of . . . engagement").

Streaks play an outsized role in causing younger users to experience anxiety and stress. Ex. 933 at 5093 (internal engineering document describing Streaks as a "self-perpetuating, anxiety driven system" with "compulsive behavior pattern[s]"); Ex. 989 ¶ 69 (same). This is because Streaks "take advantage of adolescent susceptibility to social cues by providing a key indicator that (a) an individual adolescent has social connections and (b) that they are in contact with their social connections daily….This is vitally important to adolescents in terms of their social development." Ex. 983 ¶ 87. Nonetheless, Snap designed Streaks to target its core teenage demographic, with senior personnel acknowledging it is a feature primarily used by teen users. Ex. 847 at 307:6-308:21. The feature's penetration among high school students demonstrates its effectiveness: 41% of high school Snapchat users maintain an average of 9.5 Streaks simultaneously, demanding constant attention throughout the school day to avoid losing their accumulated counts. Ex. 936 at 6711, 15. Losing a Streak can cause teenagers and adolescents enough anxiety that many contact Snapchat customer support to plead for reinstatement of lost Streaks. Ex. 929 (email to Spiegel from high school student stating, "I can't imagine how many emails you get per day from teenagers threatening you if you don't return a streak they lost"). As explained by the Districts' expert, Dr. Tim Estes, "Snap could have, and should have, turned this feature off for minors. Instead, Snap doubled down, and monetized it." Ex. 989 ¶ 72. Specifically, in 2023, Snap monetized teen stress by offering a paid option to restore a user's Streak. Ex. 937 at 0645-47 (early launch internal data predicting a "floor" of [redacted] in annual revenue from Streak restore).

Finally, Snap knew that its platform's "addictiveness" affected student performance at school. Ex. 930 at 5370. It learned this is true not only because of compulsive use during the school day, but because of the disruptive impact Snapchat has on young users' sleep, which affects student performance. Ex. 938 (internal distribution of research demonstrating negative impact of compulsive social media use by teens on sleep). After learning this, Snap conducted no further research into whether Snapchat was affecting student's academic performance and simply buried its head in the sand. Ex. 847 at 225:18-227:13.

Indeed, the company ignored repeated requests over years to assist schools in handling the unique disruptions caused by Snapchat use at school. Teachers have repeatedly contacted Snap requesting that it block the use of Snapchat during school hours or remove geofences around their school because of concerns related to student harassment or sharing of CSAM. *See, e.g.*, Ex. 979 at 7542; Ex. 853 at 376:4-379:2; Ex. 980 at 9005; Ex. 813 at 91:24-92:9, 99:17-100:21; Ex. 873 at 5839-5840; Ex. 801 at 689:22-692:25. Likewise, school administrators have complained directly to Snap's leadership that they are "slammed with Snapchat issues," and the app is a constant "distraction." Ex. 920 at 4453; Ex. 801 at 708:17-709:20. Snap knew this was true—its research observed that children use Snapchat at school to "cure boredom," Ex. 911 at 3999; Ex. 861 at 111:18-112:24, and it knows that, when bored students are distracted by Snapchat, that would "disrupt the teacher's ability to teach." Ex. 851 at 37:11-38:1. But Snap has done nothing in response to these complaints and requests.

At the 2023 National Student Safety and Security Conference, Snap's Head of Global Platform Safety, Ms. Beauchere, met with educators and school resource officers to discuss schools' experiences with Snapchat and how the app was being used in school environments. At the panel, "[a]ll participants said that Snapchat is part of their everyday life at school and said both parents and students regularly report to them harms or concerns related to Snapchat." Ex. 801 at 696:1-700:1. These educators and school resource officers described concerning interactions facilitated by Snapchat, including its ephemeral messaging, which encouraged teens to engage in risky behavior like sharing CSAM. *See* Ex. 920 at 4451 (describing parents and students informing schools about "harms or concerns connected to Snapchat" and connecting "perceived ephemerality" to, among other things, "sharing of nudes and financial sextortion of males"). One educator called Snapchat "the #1 cause of drama in school aged children." Ex. 920 at 4452; Ex. 801 at 706:7-14. Snap heard these concerns and, again, did nothing.

Given its extensive external information documenting potential harms to teenage users, as well as information it learned from user experience surveys, Snap could have and should have changed its harmful product features and, at the very least, warned users, parents, and schools. As detailed in the subsections below, it did neither of these things.

In addition, Snap made a deliberate corporate decision not to conduct meaningful internal research into whether Snapchat was harming the mental health of its adolescent users. Morgan Hammerstrom, who

leads Snap's sole user experience and product research team, testified that she was never asked to research users' experiences related to mental health. Ex. 827 at 96:22-25. Further, although Snap conducted user surveys to gauge how Snapchat lenses made all users (not just teenage users) feel, no one at Snap ever requested research into whether users experience body image issues from using Snapchat or whether the platform increases user anxiety or stress. *Id.* at 97:4-13. When directly asked whether it mattered to her as the head of Snap's user experience research if users were experiencing body image issues or anxiety related to Snapchat features, Hammerstrom testified that it did not. *Id.* at 97:14-98:5.

This research void is not incidental but systematic. Hammerstrom could not identify a single instance where mental health had been included as a question on any survey or interview that her team conducted with Snapchat users. *Id.* at 166:4-12. Out of thousands of research projects conducted or overseen during her tenure at Snap, including a period when she was the company's only research employee, Hammerstrom has never researched whether users find Snapchat addictive, the platform's impact on youth mental health, anxiety, or depression, or the impact of features like endless scroll and ephemerality (not to mention increased time spent) on teens. *Id.* at 606:3-608:7.

Most revealing, Hammerstrom testified that investigating user harm falls entirely outside of her job responsibilities. This is notable, as Hammerstrom led Snap's product research department and was involved in all research the company conducted. When asked whether she would want to know if users were addicted to Snapchat, she testified that such information "wouldn't have anything to do with [her] job." *Id.* at 151:3-9. She further stated that if a user reported Snapchat was addictive, her team would not investigate that concern but would instead focus exclusively on product-specific feedback. *See id.* at 105:7-19 (describing an example of her research as assessing reactions to whether "Snapchat buttons were purple"). To Snap's head of user research, teenage harm was simply not her concern.

Other senior Snap personnel confirmed that the company systematically ignored or avoided conducting any research on the impact of its platform on teen or adolescent (or anyone's) mental health. When asked whether Snap collects data relating to mental health impacts, David Boyle, Snap's Senior Director of Product, deferred to Hammerstrom. Ex. 830A at 174:13-175:11. He could identify no instance in which he requested any study to be conducted or data collected related to Snapchat's impact on teen mental health. *Id.* at 175:13-18. Ms. Beauchere, Snap's global head of platform safety, acknowledged that

Snap never performed research on whether any feature of the app is addictive, despite internal company conversation suggesting that both Spotlight and Streaks encouraged compulsive use. Ex. 801 at 451:9-452:5; *see* Ex. 973 at 3224-25, 44-45 (internal document identifying infinite scroll and autoplay as "unhealthy gaming mechanics" and observing that users "feel obligated" to maintain Streaks, which "become[s] stressful").

Juliet Shen, Snap's former Lead Safety Product Manager, testified that mental health was "volunteer work" at Snap because the company did not have a dedicated product manager focused on mental health and well-being. Ex. 948 at 105:11-18. Ms. Shen confirmed that she could not recall any research on whether using lenses or filters, posting stories, checking or monitoring Stories and Spotlight views, using Snap Map, maintaining and monitoring Streaks, monitoring Snapscore, using trophies or charms, or being encouraged to add friends on Quick Add impacted teen mental health. Ex. 948 at 189:6-202:12; Ex. 962. In discovery, Snap admitted it has no unit or department specifically dedicated to researching whether its features cause teen users to suffer depression, body dysmorphia, or compulsive use of Snapchat. Ex. 963 at 22-25. Spiegel, Snap's CEO, also admitted under oath that he had never commissioned research to determine whether or not, or to what degree, Snapchat was endangering youth, and acknowledged that Snap never employed in-house child psychologists, neuroscientists, or mental health professionals to advise the company on matters related to the mental health of young people. Ex. 803 at 47:8-14, 48:1-14, 50:5-13, 108:3-12.

Snap knew enough about the harmful effect of its features, their addictive potential on teenagers and adolescents, and the disruption its platform was causing in schools to merit a warning. Yet, again, Snap not only failed to do that, it attempted to remain willfully blind of the problems it was causing. Numerous high-ranking Snap personnel testified they never conducted, ordered, or recalled any analysis or in-depth research into Snapchat's impact on the academic performance of high school students or teenagers. *See, e.g.*, Ex. 847 at 226:20-227:13; Ex. 829 at 122:16-24. Nor was research conducted with teachers, school employees, or counselors to better understand Snapchat's impact on schools. Ex. 827 at 100:5-12; Ex. 948 at 330:18-23. In short, "Snap's approach is characterized by what has been described in behavioral ethics as the 'ostrich syndrome,' where one chooses to bury their head in the sand and ignore moral responsibilities and the related standards of care. The ostrich syndrome creates what scholars have

labeled as 'moral myopia,' in which one's moral responsibilities do not come clearly into focus, and it can even result in moral blindness." Ex. 993 ¶ 390.

(5) Snap's failure to exercise reasonable care

The Snapchat platform, as designed, functions as an ecosystem that encourages compulsive and harmful use by its adolescent users. Features like Snap Score, Streaks, and ephemerality, described above, gamified communication and compelled teens to obsessively monitor Snapchat to bolster their in-app persona and ensure they did not miss out on Stories, messages, and videos posted by their friends. Snapchat's Quick Add feature prompted users to recklessly accept friend requests from strangers to increase their Score, connecting unsuspecting teens with dangerous child predators. Snapchat's camera incorporated beauty lenses that drove body dysmorphia, disordered eating, and anxiety about appearance. And Snapchat adopted the addictive innovations pioneered by its social media peers, including endless algorithmic feeds, predatory notifications, and geotargeting. The sum effect of these features was to drive addiction and mental health harm among adolescents—and that was by design.

Because Snapchat relies so heavily on adolescents to drive and maintain growth, Snap has never implemented effective age verification. Snapchat's practically non-existent age gate allowed millions of users to misreport their age and access the platform before they were thirteen years old. Worse, until 2022—the year this litigation was commenced—Snapchat provided parents with ***no parental controls at all***. Even now, those "controls" are highly ineffective, not promoted by Snap, and as a result, scarcely used. As a result, Snapchat's youngest users, including millions who are not supposed to be on the platform at all, are exposed to a bevy of harmful features.

(a) Age Verification

Snapchat's process for verifying a user's age at sign-up has always been, and is, completely ineffective. Ex. 871 at 8461; Ex. 867 at 285:5-286:5 (senior director of product observing that Snapchat's registration flow resulted in "scenarios where we're highly confident the inputted age is incorrect."). Prior to 2013, Snapchat did not require users to enter their age at all during registration. Ex. 867 at 143:16-144:7. By 2015, Snap still did not know the age of approximately 25 million daily active users who were not required to submit their age at signup. Ex. 868 at 4875. When Snapchat finally introduced a birthday field in its registration flow, that field defaulted to the year 2000, a setting that automatically made users,

including those under 13, appear to be teenagers regardless of their actual age. Ex. 869B at 785:4-787:5; *see also* Ex. 1149 at 5673 (internal document stating that Snap personnel were unsure when this change was implemented). Snap modified its registration process in 2017 to default a new user's age to 18, which was no improvement since any child who failed to input their birthday would appear to be 18. Ex. 870; Ex. 863 at 118:11-14 (confirming that "if a user doesn't manually enter their real birthday they are automatically set to 18"). And while Snap fixed this issue in other countries, Snapchat in the United States to this day defaults a user's age to 18. *See* Ex. 863 at 173:17-174:1, 184:11-185:16.

At no point in its existence has Snapchat required users to do anything meaningful to verify their age when creating an account. So long as new users input a birthdate identifying themselves as 13 or older—or inputs no birthdate at all, thus defaulting to 18—they will successfully pass Snapchat's age gate. *See* Ex. 863 at 393:24-395:14 (demonstrating that Snapchat's account creation flow allows users to create an account with a fake name, birthday, and email). This frictionless age gate effectively encourages young users to misrepresent their age at sign-up—including users under the age of 13, who Snapchat's terms of service nominally prohibit from using the platform. Ex. 987 ¶ 137.

Senior Snap personnel, including its Global Head of Platform Safety Jacqueline Beauchere, acknowledge that children—especially those under 13—will misreport their age at sign-up to gain access to the platform. Ex. 801 at 482:7-484:1; Ex. 874 at 0931. Worse, when children under 13 *do* accurately report their age—and are denied registration—Snapchat permits them to immediately re-register on the same device using a different birth date. Ex. 863 at 216:16-223:15; Ex. 867 at 162:5-164:14 *See also* Ex. 874 at 0931 ("when a user tries to register and fails because they input an age under 13," they are "allowed to register again" on the "same physical device"). Snap's product teams considered prohibiting this practice and requiring "locked out" periods but ultimately opted against it. Ex. 869B at 855:6-856:13. According to Senior Vice President of Global Policy, Jennifer Stout, lockout periods "would have a negative impact on daily new users and daily active users," causing a "material impact to user acquisition and growth." Ex. 878 at 1089.

Snap does not take proactive measures to find underage accounts or remove them when they are detected. Snap has never conducted an audit to determine the number of users under thirteen that maintain Snapchat accounts. Ex. 863 at 95:7-17. The company does not track or otherwise monitor whether ages

reported on its platform are accurate. *Id.* at 251:9-16; Ex. 867 at 117:6-9. Nor has Snap conducted user research into the number of children who were under the age of 13 when they created their accounts. Ex. 876; *see* Ex. 867 at.117:6-9. Snap's public Transparency Reports do not disclose the number of under-thirteen accounts the company moderates because the numbers would be "laughable." Ex. 877 at 3323 (internal communication between product managers stating that, whereas TikTok deletes 20 million underage accounts per quarter, Snap "delete[s] like 50 per day" and obscures this data in its Transparency Reports). In December 2021 alone, Snapchat had over 43 million users identified as under 13. Ex. 875. During this same period, Snap deleted only 2,430 users—0.00005% of those identified as underage—for accessing the platform before they were 13. Ex. 874 and 0931.

Not surprisingly, Snap knows that underage users are on Snapchat during elementary and middle school, providing Snap an even earlier vector for growth. Ex. 851 at 421:7-16; Ex. 885 at 3460. Indeed, Snap's Head of User Research, Morgan Hammerstrom, acknowledged Snap was aware of children as young as the first grade using Snapchat at school. Ex. 827 at 451:5-452:20. One elementary school teacher reported that students had been creating "anonymous" Snapchat accounts, "add[ing] a bunch of friends, and then spread[ing] negative or not-appropriate content around the school," which a Snap employee described as "not unlike 'note-passing' at scale without accountability." Ex. 885 at 3460; Ex. 851 at 421:7-16. Reports such as this were shared with dozens of Snap personnel. Ex. 885. At a National Student Safety and Security Conference, school resource officers met with Ms. Beauchere, Snap's Global Head of Platform Safety, to explain that children as young as 9 and 10 years old were sharing CSAM over Snapchat during their third and fourth grade classes. Ex. 920 at 4453; Ex. 801 at 712:11-713:3. Yet still, Snap did next to nothing to proactively remove or moderate these underage accounts.

Notably, Snap has developed the capability to accurately estimate a user's age after registration—but it has only used these technologies for advertising purposes, not to detect and remove users under thirteen. *See* Ex. 807 at 161:16-162:23 (the purpose of Snap's age inference technology was "to target ads"). Snap has developed several methods to estimate a user's age based on photos and videos they post, as well as their interactions with other users and the platform. As early as 2015, Snap developed the capability to estimate user age through facial analysis. *See* Exs. 868; 879. In 2016, Snap's age inference

technology was utilized by its monetization team to display products with age restrictions, like alcohol, to appropriate age cohorts. Ex. 867 at 277:2-278:10, 279:24-280:18; Ex. 807 at 161:16-162:23.

Snap's machine learning model for its Spotlight platform can also infer when a user is under 13. In response to fines imposed on TikTok in the UK, Snap "conducted an experiment to assess the volume of potential under 13 escalations via our Spotlight Moderation workflow and escalated between 2,000-3,000 Snaps per day globally," demonstrating it has the capacity to identify and moderate under-thirteen accounts active on Spotlight. Ex. 884 at 4413. But again, that model is not used to flag underage users, only to filter content displayed to certain accounts. Ex. 881 at 0935-36 (describing the purpose of and ability to infer age below 13); Ex. 882 at 9464; Ex. 883 at 120:21-123:22 (describing the purpose of Spotlight's age inference model to filter content, not flag underage users); Ex. 883 at 98:21-99:16, 120:5-122:15 (former Head of Product explaining that "filtering out" from Spotlight means to "remove" or "make not eligible" for viewing). In other words, although Spotlight's machine learning model can identify underage users, Snap takes no action against those accounts.

Alex Osborne, Snap's former Safety Product Manager, attributed the failure to moderate underage accounts on Spotlight to "operational bandwidth limitations on the [Trust and Safety] side." Ex. 884 at 4413. But the "bandwidth limitation" was deliberate. In 2023, Osborne proposed straightforward solutions, including "simply increasing the staffing on our Trust & Safety team dedicated to age verification" or "integrating with a third-party age assurance vendor that could verify user age through selfies or ID." *Id*. at 4414. Snap rejected these proposals. Osborne explained the real concern: "there are risks with this approach. For example, regulators might use our small experiment against us and demand that we expand our usage of age estimation tools to other, more consequential parts of Snapchat." *Id*. As of November 2024—over two years after this litigation commenced—Snap had not engaged with a third-party age assurance vendor to moderate underage accounts on Spotlight. Ex. 863 at 114:21-115:16.

Snap could have utilized any one of the age estimation technologies it developed to identify users under 13 and remove them from the platform, but it has failed to do so. *Id*. 863 at 258:11-260:17. Prior to 2024, the company's sole approach to identifying and moderating such accounts was to rely on reports or complaints from other users. *See* Ex. 867 at 88:14-89:1. Even this reporting avenue was deliberately weakened, however, because Snap does not permit users to report suspected under-thirteen accounts

through the app itself; instead all reports must be made through Snap's website, adding needless friction to the reporting process. *See id.* at 249:23-261:4. Snap's seeming indifference to the presence of children under thirteen on Snapchat is particularly concerning given that Snap's founder and CEO admitted under oath that "children under the age of 13 are not ready to communicate on Snapchat." Ex. 803 at 168:6-9. The company looks the other way because children are so critical to Snapchat's growth and profitability.

(b) Parental Controls

Snap has failed to provide parents with effective controls to prevent their children from using Snapchat compulsively and/or at school. Despite Snapchat's extensive reach among teenagers and adolescents, Snap did not introduce ***any features*** resembling parental controls until 2022, eleven years after the platform was created. Ex. 801 at 259:17-260:7. At the time, Snapchat was "the only major platform that does not yet have parental controls." Ex. 853 at 409:23-421:21. Even after these tools, called "Family Center," were introduced in 2022, they were designed to be weak and serve more as a window-dressing marketing tool as opposed to an actual parental control. *See* Ex. 889 at 9334 (Family Center intended to be a marketing tool to "address parental concerns and school bans"). Ms. Beauchere instructed her colleagues that she was "wary" to refer to Family Center as parental controls "because they're not;" "traditional parental controls cover the who, what, where and how much of a young person's life," but Family Center was not designed to provide parents those resources. Ex. 890 at 9241.

Family Center allows a child to link their Snapchat account to a parent's account, after which the parent is able to observe certain actions taken by the child's account. Ex. 894 at 7950. As first introduced, Family Center permitted parents to view only the names of their child's friends and the group chats they were participating in, Ex. 1171 at 9123, but not to set time limits, restrict content, or control their child's behavior on the platform in any meaningful way. Ex. 891 ("Time limits and other forms of restrictive mediation are traditional 'parental controls.' Snap is not offering traditional parental controls[.]"); Ex. 893 at 9261 (Spiegel commenting, "I think the goal here is to give the parent visibility but not necessarily control"); *see also* Ex. 890 at 9241 (Beauchere stating, "we've gone out of our way NOT to refer to them as parental controls because they're not. Traditional parental controls cover the who, what, where, and how much of a young person's online 'life.' These are built differently."). Although certain Snap personnel considered allowing parents to control their child's privacy settings, Spiegel shot that idea down. *See* Ex.

893 at 9261, 9264; *see also* Ex. 801 at 281:15-282:-15. Snap has also declined to promote Family Center to parents or teens, *see* Ex. 896 at 0984 (Snap had no plans for "community outreach"), and "Family Center is extremely hard to find in the app," Ex. 894 at 7951. Snapchat does not prompt teenagers either at sign-up or at any point to enable Family Center protections or inform teens that Family Center exists. *See* Ex. 987 ¶ 139.

Because it is both hard to find and effectively useless, Family Center is not used by Snapchat users or their parents. *See, e.g.*, Ex. 894 at 7951-52; Ex. 801 at 303:22-304:12 (Beauchere acknowledging that the number of users utilizing Family Center is "low"). Indeed, only 0.33% of the platform's teenage users have utilized the feature. Ex. 894 at 7952. But adoption or promotion of Family Center was never the point. Instead, the primary purpose of Family Center "was to have a feature we can point regulators, interested parents, and the press to, rather than broadly tell everyone to adopt this." Ex. 897 at 6211. Why? Snap was concerned that parental controls, including even weak pseudo-controls like Family Center, may negatively impact "engagement . . . if this lands the wrong way with teens." *Id.* And Snap's overriding goal has been to maintain its popularity with teens and adolescents at all costs.

(c) Barriers to Exit

When teenage and adolescent users decide to deactivate their accounts, Snap erects barriers. *See* Ex. 898 at 2746 (explaining that Android users are unable to delete their accounts from in-app settings and need to access the web portal instead); *id.* at 2756 (iOS users not able to delete account in-app until January 2022); Ex. 899 at 4403 (inability to delete accounts on Android continued to persist into 2024). Moreover, parents face significant hurdles when trying to delete the accounts of their underage children. *See, e.g.*, Ex. 900 at 9434 (communication from parent of underage child complaining that "Your website doesn't allow parents to delete accounts").

Once a user navigates the deactivation process and is about to delete their account, they are presented with a black screen depicting a crying emoji and a message that reads, "Your account will be deactivated, which means friends won't be able to contact you on Snapchat. You'll also lose any Chats you've saved and Snaps and Chats you haven't opened." Ex. 901. The obvious goal of this message is to discourage deletion. Finally, even a user who proceeds does not actually have their account deleted—Snap

"require[es] a mandatory 30 day 'cooling off' period during which the user could disrupt and cancel the process at any time—simply by logging into the account again." Ex. 987 ¶ 151.

(d) Geolocation

Snap developed and deployed comprehensive geolocation infrastructure specifically designed to identify, track, analyze, and grow use of Snapchat at schools across the United States. The foundation of Snap's efforts was a proprietary geolocation database called "Verrazano," or "Snap's Places Database." Ex. 906 at 6440; Ex. 907. Verrazano contained approximately 16,000 places categorized as "high schools" in the U.S., with additional schools potentially categorized under "School" or "Education." Ex. 906 at 6456. Snap personnel described the purpose of this effort as creating "Snapchat world for high schools." *Id.* To enhance this effort, Snap purchased school boundary polygons from third-party vendors, which the company manually cleaned and supplemented. Ex. 907; Ex. 908. Snap also acquired granular location data, including a comprehensive list of high schools' location data, from a vendor. Ex. 807 at 138:19-140:24. This data allowed Snap's advertisers to geotarget students with a "sponsored spaces" geofilter whenever those students were within a certain radius of the school. *Id.*; *see also id.* at 137:19-138:1 (explaining how geofilter ads generally work). In 2016, Snap charged advertisers $100,000 to access only "17 high schools, which it systematically leveraged to increase ad dollars through school-by-school penetration." Ex. 910 at 0698.

Snap invested in this geolocation infrastructure so it could understand students' every action while in school and monetize that information. The company created a dashboard "to view the number of unique users, geofilter swipes, and session counts within a geofence." Ex. 907 at 3631. This dashboard enabled Snap to track real-time metrics for individual schools. *Id.* (demonstrating the tool's capabilities at an employee's former high school and showing 1,400-1,500 unique users within the high school's geofence during the school year). The analytics went beyond simple user counts. Snap calculated exact penetration rates by comparing its user data against enrolled student populations. *See id.* at 3633 (reporting that "we have 1,450 unique users out of 1,684 students for a penetration rate of 86.1%" at one high school). Using this technology, Snap was able to "measure swipe rates and users within a fence to find areas in need of geofilter," and could "easily see the student population of the high school and links to a detailed breakdown of demographics, tests, and other information about the school." *Id.*

This school-focused approach had buy-in at the executive level. In September 2018, Josh Siegel, then heading Snap Map and one of the company's top Product managers, articulated the company's growth philosophy in an email to the growth team, proposing "better ways to model growth, specifically about how we can model school-by-school penetration." Ex. 907 at 3633. Siegel's communications make clear the strategic importance Snap placed on educational institutions: "US universities and high schools are really our core, both in terms of historic growth and penetration, as well as the loudest source of user feedback for us and the primary place investors/advertisers assume we are winning. If we lose these users, we lose." Ex. 1169 at 3579 Siegel pushed for systematic monitoring of school penetration, requesting "a regular report or analysis that looks at US school penetration" and confirming that Snap possessed "all the data to figure out who these users are based on their GPS location or even age + IP-based location at the minimum." *Id.* This school-by-school growth and surveillance strategy was fundamental to Snap's expansion model. Siegel explained that understanding "school-by-school behavior and penetration" was critical for determining "how to ignite growth in new markets by focusing on school-by-school takeover tactics." *Id.*

There is no evidence in the record—anywhere—that Snap ever warned schools, administrators, and teachers that the company was collecting detailed geolocation data about their students, for the purpose of increasing their use of Snapchat during the school day. Snap's policies governing geofiltered content targeting schools contain detailed guidance regarding material likely to be disruptive to the learning environment, but nowhere do these procedures instruct employees to inform a school or its staff that their location is geofenced. *See* Ex. 848 at 5121-30, 85, 97, 55. Further, although Snap has the technology to geolocate and/or geofence schools, there is no evidence that Snap has ever used either of these capabilities to prevent access to Snapchat by students that were on school property during school hours. Ex. 829 at 328:23-329:17; Ex. 801 at 695:15-24. Indeed, Snap did just the opposite, ignoring requests from school employees who realized their schools were geofenced and requested that those geofences be eliminated. *See, e.g.*, Ex. 1147 at 6480 (accusing Snap of "falsely identif[ying] our institution's location to include surrounding neighborhoods not associated with our school or brand" and "asking one more time for [Snap] to take this more seriously"); Ex. 1148 at 9022 (explaining a geofenced Story "cannot be disabled").

(e) Notifications

The ephemeral nature of messaging on Snapchat creates urgency to quickly view, respond, and save communications, requiring users to develop daily checking habits. *See* Ex. 1154 at 2007 (ephemeral messages cause "users to develop a daily habit of checking them"). Snap recognized that this "FOMO" could be further manipulated with notification campaigns. Ex. 912 at 6474; Ex. 913. "Snapchat employs a notification recapture strategy, leveraging social prompts to draw users back onto the platform through prompts to view a friend's story, a new Snap, or a general prompt to catch up on content." Ex. 987 ¶ 140. Indeed, Snap deploys push notifications as "a powerful lever for growth and engagement," sending approximately 311 million notifications daily to drive reengagement. Ex. 914. In 2022, [redacted] notifications drove approximately [redacted] app opens daily, with [redacted] incremental daily active users attributed to growth notification pushes. *Id*.

Snapchat's push notifications are one major source of distraction and disruption in the school environment. Ex. 807 at 226:22-227:12. Snap specifically created notification campaigns targeting high school students and school-based engagement, including "an aggressive back to school growth notification plan that will include dynamic user targeting based on school start dates." Ex. 915. Snap's High School Communities program, which reached 15 million students across 24,000 US high schools, deployed targeted in-app notifications designed to drive student acquisition and activation, including messages such as "Don't miss out – your school's shared story is taking off." Ex. 916 at 2516-17. Snap scheduled High School Community notification prompts to be delivered to students at 8:00 a.m. and 2:00 p.m., to encourage students to share content to their High School Community, and internally acknowledged that this constituted a "dark patterns/nudging technique[,]" which garnered student attention and disrupted schools. Ex. 917 at 7602-03.

It was not until 2023—one year after this litigation commenced—that Snap's Product research team proposed "giving users the power to turn off notifications during school hours or when they should be working or studying." Ex. 849 at 7231; Ex. 827 at 476:4-477:21. But even after Snapchat permitted users to turn off notifications, the platform continued to prompt users to enable notifications. Ex. 830A at 256:20-258:2. Snap witnesses were unaware of any efforts made by the company, or even policies or policy considerations, to restrict push notifications that adolescent users receive when they are likely to

be in school. *See, e.g.*, Ex. 853 at 379:6-19. Snap's corporate designee testified that the company declined to implement recommendations to disable notifications for student users during typical school hours. Ex. 861 at 283:12-284:19. Further, Snap did not "specifically allow" users to "schedule quiet periods such as no notifications during school hours." Ex. 830B at 508:14-20; *see also* Ex. 1145 at 4826 (finding in late 2023 that the "main reason users disabled" all "notifications is because they want to focus on school, work, or friends when they hang out with them" but Snapchat provides no option to quiet notifications during specific time periods). Snap has presented no evidence on summary judgment that it ever warned Districts, parents, or the public about the disruptive nature of its notifications.

(f) Screen Time Management

Snap has never implemented a time management feature or user control to limit notifications during specific hours of the day; even when Snapchat provided users the option to control notifications, it was either on or off. Ex. 830B at 502:11-504:19; *see also id.* at 508:14-20 (pushing responsibility for disabling notifications onto operating systems). At one point, Snap implemented a feature allowing "quiet hours" from midnight to 7 am, during which Snapchat would not send push notifications, but scrapped the feature after it was characterized as an impediment to growth. Ex. 939 at 2437 (A/B testing tracks "[redacted] DAU increase" when feature removed). Although Snap considered "giving users the power to turn off notifications during school hours or when they should be working or studying," Ex. 849 at 7231; Ex. 827 at 476:4-16, it did not do so. Snap still does not "specifically allow" users to "schedule quiet periods such as no notifications during school hours." Ex. 830B at 508:14-20. Indeed, Snap's corporate designee testified that the company declined to implement recommendations to disable notifications for student users during typical school hours. Ex. 861 at 283:12-284:19. Further, senior Snap personnel testified that the company does not restrict the ability of teenagers to use Snapchat during the seven hours they are likely to be at school—and there is no evidence that Snap has ever taken *any* action, or implemented *any* product design features, that would impede or prohibit the use of Snapchat on school property. Ex. 829 at 326:18-327:4; Ex. 801 at 695:2-14. Nor has Snap even investigated restricting the ability of a teenager to access Snapchat during school hours. Ex. 853 at 370:20-371:11; *see also* Ex. 807 at 262:17-20 (acknowledging that Snap has never paused or blocked access to Spotlight during school hours).

Unlike certain of its co-defendants, Snap never implemented features that would encourage

teenage users to "take a break" or "timeout," or to temporarily pause their accounts. *See* Ex. 863 at 314:7-318:7 (acknowledging that Facebook, Instagram, and TikTok allow users to pause their accounts but Snapchat does not). Nor did Snap ever provide teenagers a warning that they were spending too much time on Snapchat or that they should put the app down to learn or sleep. *See* Ex. 829 at 102:23-103:12; *see also* Ex. 853 at 135:19-136:11 (acknowledging that Snap should provide warnings to educate its users and enable them to better control their experience).Snapchat users requested the ability to pause their accounts and Snap considered implementing it, Ex. 1146 at 0928, but ultimately never did. *See* Ex. 863 at 321:4-10.

(g) Algorithm and Infinite Scroll

Snapchat contains three primary algorithmic content surfaces, each engineered to maximize user engagement through sophisticated ranking systems that predict and exploit user behavior patterns. These surfaces—Stories, Discover, and Spotlight—function as distinct yet interconnected engagement mechanisms that collectively trap users in a web of compulsive content consumption.

Stories is Snapchat's ephemeral surface displaying short videos posted by a user's friends. The surface utilizes machine learning algorithms that rank friend content based on predicted engagement metrics rather than in chronological order. Ex. 1155 at 7691. The algorithmic ranking for Stories relies on user features including "age, gender, locale," and behavioral signals such as view patterns, interaction history, and predictive engagement modeling. *Id.* at 7693-95; Ex. 859 at 221:4-222:22. The Stories algorithm is paired with an auto-advance experience, which encourages users to spend as much time as possible on the platform absorbing content. Ex. 950 (describing Stories redesign to create "[a]n infinite feed of ranked, interesting stories," and "[d]rive an endless scrolling experience of interesting content"); Ex. 829 at 107:6-111:18 (describing Stories as "an experience where you can scroll endlessly"). Driving engagement on Stories is critical to Snap's business: "Increasing user engagement with the Stories tab is critical for our business," as [redacted] of Snap's ad revenue is generated from this tab." Ex. 1155 at 7960.

Discover is Snapchat's platform for professionally curated short videos featuring products or experiences promoted by advertisers or media companies. Like Stories, the surface is an endless scroll of algorithmically selected videos. *See* Ex. 808 ("We already have an endless scroll design in Discover and I think we wish it was more ~~addictive~~ compelling") (strikethrough in original). The primary purpose of

Discover, like Stories, is to increase the time a user spends in the app. *See* Ex. 1156 at 0712 (internal document identifying Discover as the "silver bullet" for engagement and observing that "Habits are built in times of boredom. Instragram LIVES in this domain"). Snap personnel likened Discover to "feed-based environment[s]" with "large and nearly 'endless' content supply," comparing the surface to "FB newsfeed, Instagram, Twitter, and TikTok"—models that "work particularly well." Ex. 954 at 9605. Indeed, Snap's engineers, when designing Discover, asked "what will most effectively create a daily habit? What best takes users down engagement rabbit holes?" Ex. 1157 at 3871.

Spotlight, implemented in 2020, is Snap's attempt to replicate TikTok's scrolling video feed. Ex. 807 at 311:21-312:1 (acknowledging that Spotlight is similar to TikTok); Ex. 1158 at 8592 (users immediately recognized Spotlight as similar to TikTok and Instagram Reels); Ex. 1159 at 4071 ("Snap Inc. admitted that Spotlight was 'inspired by other platforms,' including TikTok"). Videos on Spotlight are ranked using an algorithmic sampling model, paired with an endless scrolling and auto-advance experience. Ex. 951 at 6698 (ranking team responsible for "reducing barriers to consumption."). This algorithmic video feed dramatically increased engagement, with views increasing 65% and view time increasing more than 99%. Ex. 985 ¶ 142.

Across all three surfaces, Snapchat's algorithms optimize for behavioral proxy metrics that drive compulsive use rather than user satisfaction. *See* Ex. 1160 at 7572 (algorithm optimizes for "metrics like view time, shares, favorites, subscribes, and replies"). Each algorithmic surface also utilizes an endless scroll. Snap's senior personnel have not only characterized the endless scroll as addictive but admitted that this was Snap's goal. Ex. 953 at 3050 (senior Snap personnel observing that they wished the "endless scroll design in Discover" was more "addictive"); Ex. 954 at 9064 (internal goals for content feed: "Drive increased depth of engagement, time spent and binging behavior"). Internally, Snap personnel worried that their algorithmic feeds would further exacerbate what they acknowledged was a social media addiction crisis with teens. *See* Ex. 955 at 7920 (developing communications strategy to respond to potential question asking, "At a time when there are huge concerns about screen time—won't this fuel addiction with your young users?"). Further, despite claiming to "optimize the viewer experience beyond in-session engagement metrics," Snap admits it is "comfortable using in-session actions to proxy for longer-term effects," revealing that its algorithms are, in reality, entirely short-term and engagement-

based. Ex. 985 ¶ 147. And the effectiveness of this algorithmic manipulation has been quantified in Snapchat's own metrics showing that "Spotlight ranking [which is engagement-based] improved views and view time +65% and +99% respectively, relative to a heuristic ranking holdout." *Id.* ¶ 142. Snap further acknowledges it has "limited mechanisms for asking viewers how they feel" about their experience on the platform and lacks "an understanding of long-term viewer and creator retention, and therefore do[es] not optimize for it today." *Id*. ¶ 148.

Snap witnesses acknowledged that it should warn minor users if part of its platform posed a risk to their safety. Ex. 250 at 133:1-14. They also acknowledged that if the company was aware of a safety issue, it was obligated to educate the community and give them the tools to help users control their experience. *Id*. at 135:19-136:11. They even agreed that part of these efforts "can absolutely consist of a warning." *Id*. Yet, at no point did it warn users about the harms caused by its algorithms or infinite scroll on Snapchat.

(h) Account Recommendations

Snapchat also uses algorithms to make friend recommendations through its Quick Add feature, which launched in 2015 and was designed to maximize friending activity. Ex. 1161 at 9877 ("Quick Add contributes to [redacted] of friendship requests sent by new users, and is the primary lever for improving new user retention"); Ex. 1162 at 4435 ("As of May 2022, an average of [redacted] daily friend requests are sent via Quick Add accounting for [redacted] of daily friend requests"). The tool worked exactly as intended, prompting teenagers to accumulate as many friends as possible to boost their Snap Score and Story views, and to maintain Streaks. Ex. 1165 at 3854 (describing "unhealthy streakers" as maintaining "streaks with a lot of random people . . . from Quick add"). By 2021, Snap personnel confirmed teenagers "actively [choose] to keep their privacy settings open, accepting strangers' quick adds and participating in shout outs and streaks in order to gain more contacts and 'views' on Snapchat and eventually increase their 'score.'" Ex. 806 at 6130. Internal documents recognized that teens' "Main motivation to having a lot of friends is to increase Story Views and Snapscore," meaning that "[u]sers would add a bunch of people from Quick-add." Ex. 1164 at 1318.

Unfortunately, Quick Add became a vehicle for child predation. Law enforcement agencies globally reported that over 80% of child sexual grooming cases were associated with Snapchat. Ex. 1166

at 8468. When advocacy groups created test accounts as 13-year-olds, they "immediately got recommendations for contacts in quick add who were strangers." Ex. 801 at 358:23-359:1. User reports from October 2020 flagged concerns about "suggesting minors as friends in 'Quick Add.'" Ex. 830B at 616:10-617:1. As early as 2018, Snap knew that "grooming on Snapchat" occurred because of "gamifying snap score" to add friends through Quick Add. Ex. 1167 at 6144987. Yet six years after launch, in December 2021, Quick Add's algorithm still did "not utilize any age-related feature," freely recommending adults to minors and vice versa. Ex. 956 at 7701.

In 2022, Snap purportedly attempted to alter the Quick Add algorithm to make it safer but those changes proved largely ineffective. A 2023 study by Common Sense Media surveyed girls aged 11 to 15 and found that 57% had been contacted by strangers on Snapchat in ways that made them feel uncomfortable. Ex. 1144 at 1228. When presented with this study, Snap's director of strategy testified that this issue came up "again and again" in safety conversations regarding Quick Add. Ex. 853 at 300:12-301:23. The fundamental problem is that Quick Add cannot function safely when Snap lacks reliable information about users' actual ages. *See* Ex. 1170 at 8196-98 (Quick Add protections "rel[y] on knowing the age of the user reliably" and citing an instance in which a child was sexually groomed through Quick Add because she reported her age as 22). And, as set forth above, Snap has taken insufficient steps to ensure it has reliable information about its users' actual ages.

(i) Beauty Filters

Snapchat implemented lenses in 2015, which allow users to smooth and whiten their skin, enlarge their eyes, and incorporate other beauty enhancements on photos taken by a user. Ex. 940 at 0990-91. Beginning in 2018, if not earlier, Snap possessed extensive evidence that its beautification filters contributed to body dysmorphia and eating disorders among young users. *See, e.g.*, Ex. 941 (Journal of American Medical Association ("JAMA") article discussing the impact of edited selfies and photo filters on perceptions of beauty and body image, which it called "Snapchat Dysmorphia"); Ex. 942 (internal emails reacting to the JAMA article, with Snap personnel stating, "the idea that filters have given us unrealistic expectations of beauty is well-traveled territory"). Internal communications acknowledged Snapchat users developing "Snapchat dysmorphia," a specific form of body dysmorphic disorder linked to the way in which Snapchat's beauty filters alter the appearance of teenage girls. Ex. 829 at 190:1-

191:13; Ex. 942. Snap's head of user research, Ms. Hammerstrom, acknowledged that lenses can cause "major and emotional pain" to users. Ex. 945 at 2854.

Snap could have done the right thing and addressed this concern. Instead, it focused on counter-messaging. Ex. 944 (2018 emails discussing the JAMA article on Snapchat dysmorphia and stating that "continuing to not engage w/ media" is ineffective, and that Snap must push a narrative that lenses "aren't just about beautification. They are silly and wacky"); Ex. 943 (Snap must use the "opportunity" to "find ways to push hard on a counter narrative"). When Snapchat finally—and belatedly—conducted its own research in 2020, it obtained comprehensive data that many users found lenses "troubling" and personally experienced "Lens Dysmorphia." Ex. 946 at 2481, 89, 95-96 ("I look whitewashed," "Lenses aren't made for people of color," "the filters actually make me look decent instead of being a horrendous black girl," "makes my face look thinner and lighter," "I love the filters it hides my ugliness [sic]"). Young users overwhelmingly listed beautification filters as the reason they felt they looked best in the Snapchat camera. *Id.* at 2495-96. Snap's research concluded that "the prevailing mental model among [younger users] suggests they don't know they are supposed to look any other way [than] in the Camera or in a Lens." Ex. 945 at 2854. Snap personnel acknowledged, "using a Lens can create or enhance the dysmorphia users are trying to avoid." Ex. 947 at 6231. Snap Product Manager Juliet Shen testified that she identified "articles or other types of studies showing that [skin lightening filters] could be harmful" and raised these concerns with colleagues at Snap, but the filters still exist on the platform. Ex. 948 at 221:4-222:16.

Snapchat still pushes beautification filters on its teenage users to this day. Why? Because Snap understands that beautification filters substantially increase engagement. Ex. 949 at 6749 (Snap's Senior Product Manager for Lenses stating beautification lenses "bri[n]gs us +70% of engagement"). Snap has never warned parents or their children that beauty filters contribute to problematic use, social comparison, eating disorders, depression, and/or anxiety. *See* Ex. 997 ¶¶ 150, 185, 189, 192, 195, 199, 202; Ex. 983 ¶¶ 120-21.

(j) CSAM Reporting

In the Court's Order resolving Defendants' Rule 12 motions, it made clear that each platform's processes for allowing users to report suspected CSAM was not barred by Section 230 or the First Amendment. *See* MTD Order, 754 F. Supp. 3d at 963. Snapchat was (and is) a cesspool of pornographic

or semi-pornographic content, as even Meta CEO Mark Zuckerberg has privately recognized. Ex. 304 at 2387 ("I just logged into Snapchat for the first time in a while…What surprised me is how porny and inappropriate the content they're featuring is in their equivalent of Reels (Spotlight)…I was completely blown away that they haven't faced more scrutiny for this, especially given that their audience skews so much younger."). Much of this content involves minors, and some of it is child sexual abuse material ("CSAM"), which is illegal contraband under federal law. *See, e.g.*, Ex. 801 at 709:21-710:16 (explaining that teens and adolescents exchange nude photos and videos in group chats and Stories); Ex. 920 at 4453 ("sharing of nudes was described as the 'never-ending' problem"); Ex. 993 ¶ 404 (quoting Snap internal documents and observing that "20% of girls and 13% of boys aged 14-18 said they had shared a nude image or video" on Snapchat); *see also* Ex. 853 at 282:21-283:10 (explaining that Snap does not remove pornography from the platform unless it is reported).

Sadly, Snap has failed to provide users adequate tools to report CSAM. For most of its existence, Snapchat did not have a dedicated option to report CSAM. *See* Ex. 1168 at 19 (as of 2020, no CSAM reporting options for Snaps or Stories). Users wishing to report CSAM were required to identify it through some other reporting category. *Id.* (users may report nudity or sexual content but no option to flag the involvement of a child). Additionally, Snapchat had no options to report a user for sharing CSAM. *Id.* Even more egregiously, prior to 2023, Snapchat lacked any mechanism through which users could report content shared through its chat function, despite chat having launched in 2014 and even though Snap knew that the most severe forms of child exploitation occurred through chat. Ex. 1066 at 160:18-162:22; *see also id.* at 167:14-169:22 (acknowledging that most child exploitation occurs in the chat feature); Ex. 960 at 6319 (internal report stating, "TLDR; while chat media is reportable, chat text is currently not. We get feedback from customers, parents, legislators, and advocacy groups that abuse happens in chat text that goes unnoticed due to the lack of reporting functionality"). Even after Snap belatedly launched chat reporting, the volumes of new workflows overwhelmed its Trust & Safety operations and the team was unable to address most reports. Ex. 1131 at 9216. In addition, though Snap personnel recognized that child grooming is "the hardest problem on the platform, particularly given the app's ephemerality, Snap did not implement a proactive grooming detection framework until 2022. Ex. 948 at 260:20-261:4.

Snap does not review the reports of CSAM it receives. Snapchat's Transparency Report misleadingly states that virtually all reports seeking content removal are resolved within a reasonable time. Ex. 958 at 147:14-150:15; Ex. 1132 at 7225. But it fails to mention that this statistic does ***not*** include reports Snap didn't review before the underlying content expired (or disappeared automatically). Ex. 958 at 147:14-150:15. Expired, and therefore unreviewed, reports included approximately 46% of reports for sexual content, 20.5% of reports for harassment, and 17.3% of reports involving threats of harm or violence. *See* Ex. 961 at 3867-68. Again, because for most of Snapchat's existence there was no option to report CSAM as such, it stands to reason that these expired reports included reports concerning that illegal content. Further, Snap almost completely failed to review reports of *accounts*, including those trafficking in CSAM. Snap personnel testified that for over four years, users were able to report accounts for violations of the terms of service, but these reports were not reviewed. Ex. 948 at 340:18-342:1 (explaining the lack of workflow available to review all or even a majority of the account reports); Ex. 958 at 154:14-17 (agreeing that Snapchat was "incurring substantial risk from not reviewing these reports"). In 2022, ***96.7%*** of account reports were left unreviewed. Ex. 957 at 2520 (note explaining "significant gap" in this "critical safety tool" and that "the content reporting menu no longer reflects [Snapchat's] internal policies" or what Snap "report[s] out on the Transparency Report").

(6) Snap's misrepresentations and manipulations

As detailed above, Snap went out of its way not to research issues concerning teen mental health—in order to maintain a posture of plausible deniability. But to be clear, Snap did not remain completely passive either. What it chose to do—where it elected to put its resources—was in developing counter-narratives to dispute, minimize, or reframe evidence of harm. As early as January 2017, Snap was developing "Helpful Pushback Narratives" in response to press articles describing Streaks as a "digital obsession." Ex. 964 at 8501-05. These narratives were crafted to counter reporting that quoted therapists detailing that "the makers built into the app a system so you have to check constantly or risk missing out," and describing Streaks and Snapscore as "unethical designs [which] exploit psychological vulnerabilities to influence what users do without realizing it." *Id.* Rather than rigorously determine if that were true, Snap focused on insisting it wasn't.

In August 2018, when the Journal of the American Medical Association published an article on

body dysmorphia and Snapchat, Snap immediately developed "Sales Messaging" to be "used reactively with clients," downplaying "anecdotal evidence from plastic surgeons" and instructing all employees "speaking at a conference, on a panel, or large event about AR" to review the messaging so they could stay on point. Ex. 965. Internal communications reveal the company's dismissive attitude toward the emerging evidence. Snap's senior communications director described "Snapchat Dysmorphia" as "clearly a made up trend story" and complained that "now it's in a JAMA journal and now people will cite JAMA as concrete-sounding evidence for Snapchat dysmorphia." Ex. 966 at 5654. The concern was not whether beautification lenses were causing mental health harm—Snap made no effort to research that issue for two more years (at which point it found out the article was right). *See* Ex. 827 at 372:13-23, 562:4-9 (acknowledging Snap conducted a user survey on lenses but it was not intended to investigate mental health concerns). The concern was whether the term "Snapchat dysmorphia" would damage the company's reputation.

Snap's strategy extended beyond reactive messaging to proactive reputation management. Internal documents show that Snap created what it called a safety cross-functional team—but rather than prioritize safety, it focused on "PR and communications for parental perception" over mental health research and interventions. Ex. 967 at 1617; Ex. 830A at 143:6-160:16. When Snap undertook research on parental perceptions of its platform, the express focus was on developing hypotheses that would "point to opportunities for the company to meaningfully improve parents' perception of the platform" rather than implementing safety measures that parents were actually concerned about. Ex. 968 at 9094-9096; Ex. 803 at 175:1-177:15. This strategic approach, which Snap internally referred to as "Project Butterfly," involved coordinating marketing, communications, and public relations efforts to manage stakeholder concerns while systematically avoiding meaningful research into whether the platform was causing documented harms and addressing the underlying cause of those harms. *See* Ex. 830A at 184:17-186:11.

In March 2019, Snap compiled a document cataloguing negative press coverage about Instagram's impact on teen mental health and concluded that there was "a slight advantage for Snapchat." Ex. 969 at 3846. The document noted that "Instagram was ranked as the worst offender of young people's mental health and well-being" while "Snapchat was ranked second behind Instagram," and quoted studies finding that those who limited their use of social media apps "showed significant reductions in loneliness and

depression." *Id.* The clear implication was that if Instagram faced greater scrutiny, Snapchat might escape attention, despite Snapchat's own ranking as the second-worst platform for youth mental health. A responsible company would have taken the message that it needed to mitigate the clear negative impact its platform was having on its youngest users. Snapchat did nothing of the sort.

This emphasis on narrative management extended to Snap's most senior leadership. In October 2019, when CEO Evan Spiegel gave an interview about Streaks and mental health, he stated that "we've studied this obviously extensively," and claimed that using Snapchat "every day is actually very enriching and it's very different than this idea of community competing." Ex. 970. These statements were made despite the fact that Snap had conducted no comprehensive internal research on mental health impacts and Morgan Hammerstrom, Snap's head of user research, had never been asked to study whether the platform affected users' mental health. Ex. 827 at 96:22-97:13.

By 2021, Snap had institutionalized this approach through a comprehensive "message book" designed "to provide messaging on key issues for use in media engagements." Ex. 972 at 1370. The message book explicitly instructed employees that "[e]very question is an opportunity to: Reinforce the 'camera company' definition by repeating the word 'camera' and the camera messaging." *Id.* This strategic framing allowed Snap to deflect questions about social media harms by repositioning itself as a technology company focused on visual communication rather than a social media platform subject to the same concerns as its competitors. This is precisely what Snap seeks to do here. *See, e.g.*, ECF No. 238 at 1 (Snap counsel characterizing platform as a camera application different from its co-defendants and "primarily used for direct communication between people who already know each other in real life"). Whether or not that is an effective public relations strategy, it fails to stand up to scrutiny and is certainly no excuse for ignoring the very real harms that Snapchat was causing schools and students across the U.S.

(7) Snap failed to warn parents, educators, and schools about known mental health risks

Spiegel acknowledged under oath that Snap has "an enormous responsibility to keep [the Snapchat] community safe," that "protecting [Snapchat users] is Snap's moral responsibility," that the "safety of minors on Snap is [the company's] highest priority," and that Snap owes teenagers on Snapchat "a heightened standard of care." Ex. 803 at 16:13-18:5, 26:21-28:3. Further, Spiegel agreed that companies

should warn users of known safety risks on a technology platform and take all reasonable measures to protect teenagers on those platforms. *Id.* at 20:3-21:5. These sentiments were echoed by Snap's senior management, acknowledging that Snapchat's core audience is vulnerable and it therefore has a heightened duty to protect them. *See, e.g.*, Ex. 801 at 148:21-149:3, 521:1-5. Snap's Senior Director of Product Strategy, Nona Yadegar, explained that the company has "a responsibility to educate our community about the risks . . . they can experience using Snapchat," and one way to accomplish this is by providing proper warnings. Ex. 853 at 133:2-14; *see id.* at 124:23-125:2.

Despite all this, Snap witnesses testified that the company has never warned parents about many known dangers that exist on the platform. Ex. 948 at 281:7-282:8. Nor has Snap ever provided a warning to school districts or teachers that their students were being targeted with advertisements while at school. Ex. 807 at 230:7-231:8. As early as November 2012, one of Snap's outside consultants recommended that the company partner with the Parent Teacher Association ("PTA") to educate Snapchat's user base on safety, security, and privacy issues. Ex. 974 at 8656 (presenting to Spiegel on how to "[p]osition Snapchat as a leader in Internet safety"). Snap did not act on this recommendation or use the PTA to warn the public about known harms and risks.

Nine years later, in 2021, Snap communicated with PTAs to jointly survey parents about their knowledge regarding safety features on Snapchat. It was clear during this outreach that parents were interested in learning more. *See* Ex. 975 (explaining that the surveyed parents "didn't know about any" of Snapchat's purported safety features, "but the good news is that they want to help us get the word out and spread awareness to the PTA community"). At least one PTA expressed interest in working with Snap. Ex. 976 at 9617. Yet nothing progressed for the next three years.

In 2024, Snap met with the National PTA and agreed to work together to provide resources to parents. Ex. 977. At this time, the representative from the National PTA stated, "it's just bananas to me that we aren't collaborating with Snap on engaging and educating parents, all things considered." *Id*. at 2531. Unfortunately, Snap's goal in collaborating with the National PTA was not to actually warn parents, students, and schools about the risk of Snapchat use—which it has not done to this day. Instead, Snap's goal was to foster additional growth, as it wished to "create opportunities for kids to share what they love about snapchat with their parents/get parents on Snapchat to understand/use/play with features." *Id.* Even

this self-serving opportunity ultimately fizzled, after Snap's Policy team cited the "questionable" return on investment. Ex. 978 at 2670 (Beauchere email stating that "the ROI was questionable").

Snap worked with other school organizations to create a Guide to Snapchat for educators in 2024, but it failed to include in this Guide any risks associated with Snapchat use in schools or by students. Snap personnel described their goal to create a better experience within school as it relates to Snapchat. Ex. 801 at 714:2-25. Although this Guide includes issue-specific resources, it does not warn about the addictiveness of the Snapchat platform or its tendency cause mental health issues, disrupt the classroom, and harm the school environment. *See* Ex. 1043.

As Plaintiffs' expert, Dr. Minnette Drumwright explains, "Snap's actions are a prime example of pseudo corporate social responsibility. Snap 1) ignored and failed to disclose the harmful and negative experiences that youth encounter on its platform, 2) ignored and failed to disclose that it had inadequate age management technology and processes, and 3) made product decisions that exacerbated addiction among youth and failed to disclose or warn of the dangers. Moreover, in assuring teens, parents, the US Senate, and the American people that Snapchat was safe for teens, Snap did the opposite of warning—it falsely reassured people that Snapchat was safe for young users." Ex. 993 ¶ 388.

d) Google/YouTube's Conduct

(1) YouTube's business model

YouTube LLC, a subsidiary of Google LLC (together "YouTube"), is a multibillion-dollar company that profits from user engagement and on-platform advertising. *See generally* Ex. 700. YouTube makes money through two primary channels: advertising revenue—which necessitates an emphasis on increased watch time—and paid subscriptions. *See e.g.*, Ex. 701 at 1287 (specific product goals tied to "aim of driving paid subscribers"). YouTube has an expressed goal of fulfilling its financial targets by reaching every child in the world every day. Ex. 702 at 154:25-160:24; Ex. 703 at 3. And YouTube is good at doing just that—it boasted about creating the "YouTube Generation," young people who prefer to binge watch YouTube (for 4+ hours a day) instead of watching cable television. *See* Ex. 704 at 34.

YouTube's ability to leverage users to turn a profit is undeniable and astronomical. YouTube's 2015 goal was to reach *one billion hours* watched *a day*—expecting this would translate to $15 billion in revenue at then-current monetization rates. Ex. 705 at 2, 5. Since then, its profit margins have steadily

increased. In 2020, YouTube's advertising revenue was $19 billion, Ex. 707 at 2522; by 2023, it had skyrocketed to $31 billion. Ex. 700 at 5637; *see also* Ex. 700 at 5609; Ex. 706 at 2396.YouTube is consistently touted as a "primary" generator of advertising revenue within the Google and Alphabet pantheon. *See* Ex. 700 at 5630; Ex. 708 at 3081; Ex. 706 at 2416; Ex. 707 at 2523; Ex. 709 at 2655.

YouTube's success is built not only on its number of users, but also their frequency of use. The better it can monopolize and hold user attention, the more YouTube stands to gain. Accordingly, YouTube has prioritized user "engagement" and "Watch Time" as goals. Ex. 710 at 3; Ex. 711 at 3. "[E]ngagement" means nothing more than "quantity of consumption" (and increasing engagement just means "driving greater consumption of videos"). Ex. 772 at 7889. Likewise, Watch Time is straightforwardly defined as "How much time the user spends watching videos." Ex. 713 at 1686. Watch Time "has a long and famous history at YouTube" as a key metric. *Id.*; *see* Ex. 710 at 3 (2017 growth team presentation acknowledging that total watch time has always been a key metric); Ex. 712 at 4105 ("primary metric that YouTube's Search and Discovery systems optimizes around is time"); Ex. 985 ¶ 130.

YouTube's goal of increasing revenue by driving more engagement and Watch Time resulted in two areas of focus, detailed below. First, YouTube targeted young users—the younger the user, the longer the potential lifespan of engagement and, in turn, revenue stream. YouTube specifically targeted educational spaces, seeing an opportunity to engage a captive audience, build early brand loyalty, onboard kids into the broader Google ecosystem, and dominate young lives during their likely first interactions with technology outside the home. *See* Ex. 740 at 148:13-149:5; Ex. 741 at 10, 18. Second, YouTube designed for addiction—knowing that sustaining user engagement and attention required features that would cause users to return more and more often. Among other design choices, YouTube aims to fill even small moments in young people's lives with watch time, using push notifications to deliver "snackable" content (30-second to minute long short-form videos) for rapid consumption, while users wait for the bus or stand in line. Ex. 715 at 380:9-19; *id*. at 523:16-524:5. YouTube did this without any warning to the public or school districts—despite knowing that "[d]riving more frequent daily usage [was] not well-aligned with...efforts to improve digital well-being." Ex. 715 at 389:9-20; Ex. 772 at 7900.

(2) YouTube's targeting of school-aged children

According to its witnesses, YouTube is not a social media platform at all, Ex. 1012 at 232:7-9, but

a mere "video distribution platform." Ex. 1003 at 200:25-201:3; *see* Ex. 1001 at 280:14-17 ("[Y]ouTube is not a social media app. We don't have a social graph."); Ex. 1002 at 330:1-10; 343:11-13. Regardless of what it is called, the reality is that YouTube competes fiercely with its co-defendants in the race to attract and hold the attention of young users. *See* Ex. 772 at 7898 (discussing competition with Facebook and Instagram's feed-based consumption model which allows for ad scaling and revenue growth)) Ex. 701 at 1275, 77 (comparing design features of Snapchat and Instagram to YouTube). And, by many measures, YouTube is winning that race. Through a combination of ease of access, addictive features, and embedding in the school environment, YouTube has become the most popular platform for teens. A 2024 Pew Research Center study concluded that "73% of teens say they go on YouTube daily, making YouTube the most widely used *and* visited platform we asked about. This share includes 15% who describe their use as 'almost constant.'" Ex. 1038 at 2 (2024 Pew study comparing social media platforms YouTube, TikTok, Instagram, and Snapchat (emphasis in original)).

YouTube has dedicated significant resources to growing its youth user base, which it knew could be a rich and enduring source of user engagement and profit. "If we want to be successful with teens, besides building features to retain them once they arrive on [YouTube], we also need to focus on the experience for the next generation of teens—today's younger kids and tweens." Ex. 724 at 5689. "This is an existential question for [YouTube]—We need to make big bold bets in this space if we want to ensure that we keep these users on the platform to become teens on [YouTube]." *Id.*; *see also* Ex. 725 at 13 (Investment in youth is a "business opportunity for YouTube" to "retain and grow the next generation of creators, otherwise they could leave for other platforms and never come back"); *see also* Ex. 726 at 9331 ("Aging-up is of strategic importance to YouTube overall…Thus, defining a cogent strategy to address the 8-13 market is a key challenge and opportunity for YouTube").

To that end, YouTube developed a suite of products that would indoctrinate children early into the YouTube experience and create a seamless product progression that follows them as they grow up. In 2015, YouTube launched YouTube Kids, which it designed to attract children under the age of 13. *See* Ex. 730 at 1; Ex. 703 at 3-4 ("YTK Vision: Create the next generation of YouTube Users" and make "YouTube (Kids) as a part of the daily routine for all kids under 13") (emphasis in original); Ex. 727 at 7 (target audience is under 13 year olds). In designing YouTube Kids, YouTube identified sub-age groups

to target with specific content, with categories of users ages 8-12, 5-7, and even 4 and under. *Id*. As one executive succinctly framed it: "is there strategic value to having a kids product at [YouTube][?]…yes…they grow up[.]" Ex. 728 at 1442; *id.* at 1443 ("we have [YouTube] kids for them to shed as their 'fisher price' moment"). YouTube Kids was thus understood to be a gateway to use of the main YouTube platform.

Despite targeting YouTube Kids at children as young as four years old, YouTube made seemingly little effort to understand the developmental impacts of this platform. Shimrit Ben-Yair, the Project Manager for YouTube Kids at its inception, denied having discussions with experts regarding children and screen time (aside from speaking to her own children's pediatrician, who advised that all screen time be limited to two hours a day). Ex. 1051 at 248:12-25. To the contrary, Ben-Yair admitted that "a lot of our initial work on YouTube Kids was guided by our own experience as a parent, and as parents," *id*. at 106:3-106:21—not exactly the kind of rigorous diligence one would expect from a Fortune 10 company. Indeed, this slapdash, anecdotal approach led to design decisions that actually *contradicted* what little research YouTube undertook. YouTube Kids was initially available only on tablets. *Id*. at 76:16-24.,105:2-10. Ben-Yair testified that this decision was guided by hearing that tablets were families' preferred method to gather and watch videos. *Id*. In actuality, YouTube's official research found that television is still most often used for kids 5 to 7, bigger screens are better for parental "earshot monitoring," and educational media time drops by 16 percent when kids are on mobile devices (perhaps because "young kids are often holding the tablet" when it is made available). *Id*. at 96:9-106:2; Ex. 1006 at 5-6, 11, 43-45.

Not surprisingly given its origin story, YouTube Kids was and is ineffective as a supervision tool. First, as detailed more below, any child can continue to access the main YouTube platform, either without an account or with one (which is easy for any child to create). *See* § III.A.2.d.5.d. Second, the supervision aspects of YouTube Kids were woefully deficient in terms of addressing addictive and compulsive use. Parental controls for YouTube Kids centered on content moderation—limiting the corpus of content available, blocking certain content, and disabling search. *See* Ex. 730 at 3. While YouTube Kids had a built-in timer to set screen time limits (*see id.*), it also retained many addictive features, like autoplay, notifications, recommendation systems, and scrolling.

In fact, when YouTube launched autoplay on YouTube Kids it did so without a "toggle," i.e., without any ability for anyone—parent or child—to turn it off. *See* Ex. 789 at 1972 (2019 document discussing the *plan* to add an on/off toggle). Ben-Yair's only justification for this design decision was that, "if you as a parent said that you want your kids to watch 15 minutes of YouTube, but then every two minutes you had to go play the next movie for them, then that would not be the experience that parents wanted." Ex. 1051 at 91:5-16. In reality, what YouTube's research showed is that parents wanted *control*—the choice to make decisions for themselves and their children, including whether or not such a feature should be enabled. Ex. 1006 at 5-6, 11, 43-45. It was not until 2022—*seven years* after its launch—that YouTube created a toggle (and set autoplay to default off). *See* Ex. 1031 at 38:1-39:2.

YouTube Kids was not the company's only effort to target an under 13 year old audience. After its launch, YouTube recognized that it had a product gap for tweens—and needed a way to retain those users, or risk losing them to other platforms. *See* Ex. 731 at 6490 (raising the alarm that "**[t]here is no path today for kids to participate/create on YT**…users are trying to satisfy this demand elsewhere," citing TikTok as an example (emphasis in original)). So, in 2021, YouTube launched SupeX, an experience designed for tweens that was intended as a bridge between YouTube Kids and YouTube Main. *See* Ex. 729 at 6 (discussing YouTube's retention gaps); *see also* Ex. 730 at 3 (SupeX "designed for tweens transitioning to YouTube Main"); Ex. 731 at 6489 ("SupeX has the content missing from [YouTube Kids]…and would satisfy this age group"). Like YouTube Kids, SupeX is deficient as a supervision tool, as discussed in more detail below. *See infra* § III.A.2.d.h.

In the hunt for young users, YouTube also began to emulate features of other social media platforms. By 2018, YouTube determined that it was "the most widely used internet Platform by US teens (who watch an average of 1 hour and 22 min / day)." Ex. 772 at 7914. But this wasn't enough—YouTube was "behind Snapchat in the peer-to-peer space and Instagram in the public short form video space," so it scrambled to figure out what features it could offer to remain competitive. Ex. 701 at 1275, 77. In the hunt for young users, YouTube began to emulate features of other social media platforms. YouTube launched "YouTube Shorts" in 2020, which mimicked the vertical, short-form video format of TikTok's For You Page and Instagram Reels. *See* Ex. 1052 at 229:20-230:10. YouTube did so because it knew that this format was popular with young users. *See* Ex. 733 at 8317 ("Shorts is our big thing for teen appeal"). In

tandem with Shorts, YouTube employed notifications to maximize Shorts visits, sending users messages about Shorts at what it perceived to be moments ripe for "snackable consumption," i.e., watching content that can be viewed in little bits. Ex. 715 at 380:9-19; *id*. at 523:16-524:5.

YouTube also sought to capitalize on the popularity of effects/filters, including "Appearance effects" (defined as "all effects that alter someone's appearance"). Ex. 1071 at 6. YouTube knew these features resonated with young users on competitor social media platforms. *See* Ex. 734 at 6752 (exploring opportunity to engage users with effects and noting that Snap, Meta, and TikTok have over 2 million effects with over 100 million daily active users). It also knew that "expert researchers found beauty filters and negative social comparison may be correlated with negative wellbeing, especially in youth" and "that the pervasive nature of filtered images regularly trigger body dysmorphia." Ex. 1074 at 7.

(3) YouTube's targeting of schools

YouTube knew that it was changing the behavior of a generation—in which 53% of teens said they "would drop what they're doing to watch YouTube," 68% were aspiring influencers, and 45% thought it was important to define their online personalities. Ex. 704 at 28, 41. But even this wasn't enough. YouTube wanted the attention of young people during the school day. For more than a decade, YouTube has relentlessly pursued school adoption and student use during the school day, launching numerous platforms and tools aimed at increasing YouTube's school presence and encouraging use by students outside of the classroom. By 2013, YouTube was examining the "smartphone / tablet / app explosion" in "Pre-K / Primary," and investigating the possibility of "flip[ping] the classroom" in high school, *see* Ex. 752 at 24—by which teachers would replace in-class lectures with YouTube videos assigned as homework. Ex. 752 at 24-25.

YouTube's widespread use in schools is no accident. Google seeks to get kids using the Google "ecosystem"—apps and devices—early and extensively. Ex. 741 at 9, 14, 16; Ex. 1135 at 1, 2, 7, 11. Kids are a "Massive Opportunity" for Google because early use builds "brand trust and loyalty" that impacts kids' choices throughout their "lifetime"—particularly as to which apps to use. Ex. 741 at 1, 5, 9, 16, 17. Google quickly realized that the key to unlocking this opportunity is getting kids using the Google ecosystem at school. *Id*. at 17 ("Investing in schools helps onboard kids into Google's ecosystem."). YouTube saw two key benefits from infiltrating schools. First, doing so would increase profits via

increased watch time of educational content (i.e., increased ad revenue) and/or increased paid subscriptions for its education-tailored services. Ex. 742 at 4; Ex. 752 at 9, 17. Second, infiltrating schools would establish a user "pipeline" to ensure the youngest generation was brand-loyal to as many Google products as possible (Chromebooks, G-Suite, YouTube, etc.). Ex. 740 at 148:13-149:5; Ex. 741 at 10, 18 (explaining that onboarding kids leads to brand trust and loyalty over lifetime and school laptop brand influences future purchase patterns); Ex. 742 at 4 (noting the opportunity to create a pipeline of future users and creators); Ex. 740 at 193:2-195:11.

Google's focus on kids in schools was unprecedented. Using a fit-for-purpose "K-12 Sales Team," Google "took over" K-12 learning in the U.S. in about five years, pushing aside other tech giants like Apple and Microsoft. Ex. 741 at 16 (citing "Google Took Over Classroom" case study), *id*. at 49 ("Google has won the K-12 education technology market"); *see generally* Ex. 1136 ("North America K-12 EDU Sales Team Overview"). The key to Google's "ecosystem play" in U.S. schools were its low-cost Google Chromebook and its free apps for teachers and kids to use—including YouTube. Ex. 741 at 17-18 ("Ecosystem play" based on which laptop brand used and "[e]ven without Google hardware, kids are still using Google services."). Chromebooks could be used by kids in classrooms to watch videos on YouTube, Ex. 740 at 158:14-24, and YouTube in school was another way to embed kids into the Google ecosystem. *Id*. at 148:13-149:5. Indeed, "video" was one of the pillars of Google's effort to "win in the EDU space" and the K-12 Sales Team saw itself as providing a "bridg[e]" to "all of Google," including "YouTube." Ex. 1135 at 2; Ex. 1136 at 2.

Having inserted itself into the classroom environment, YouTube knew that it needed some way to normalize that state of affairs to teachers and principals. Over the course of more than a decade, YouTube has rolled out various palliations, which (until well after these lawsuits were filed) have been largely ineffective and designed as much with YouTube's business objectives in mind. These platforms and tools include: (1) YouTube for Schools/EDU-mode (2011), Ex. 743 at 4725; Managed Restricted Mode ("MRM") (2015), Ex. 745 at 25; Ex. 744 at 9512; Player for Learning (conceptualized in 2020 but launched as Player for Education in 2023), Ex. 746 at 7; Ex. 747 at 11; Ex. 749 at 2038; and school time controls (2024), Ex. 1002 at 205:15-206:10; Ex. 766 at 470:2-20.

YouTube's first attempt to make its platform more palatable to schools launched in October 2011 and allowed schools to sign up for what it called YouTube for Schools. Ex. 1076 at 9257; Ex. 743 at 4725. YouTube for Schools, which launched the platform in what YouTube called "EDU-mode," was a selection of videos YouTube determined were "appropriate for educators." Ex. 743 at 4725; Ex. 740 at 5-15. EDU-mode restricted which videos could be played and disabled other functionalities. Ex. 743 at 4725. YouTube also funded EDU channels, which are channels offering academic videos related to subjects taught in school or videos related to digital, professional, and trade skills. Ex. 1077 at 4; Ex. 752 at 12.

Internally, YouTube was not coy about its motivations for prioritizing EDU-mode. YouTube imagined a world where "Parents ask their children 'Why aren't you watching more YouTube?'" and "School Administrators shift budgets from Textbooks to YouTube subscriptions." Ex. 752 at 10. It understood that "EDU has [the] opportunity to be [a] key driver of view growth" and that "More EDU content = More watch time per user." Ex. 752 at 9, 17; Ex. 740 at 126:10-131:24; *see also generally* Ex. 753 (2015 internal memo identifying "YouTube EDU Opportunities"). Indeed, in the aggregate, YouTube's EDU channels garnered over 3.5 million hours (400 years) of watch time. *See* Ex. 752 at 13.

Goals aside, YouTube for Schools was known by YouTube to be a complex technical solution that was easily circumvented by students. Ex. 743 at 4725. The solution required schools to sign up and then configure their network to send all outbound HTTP traffic to YouTube with an HTTP header or URL parameter identifying their school. Ex. 1076 at 9257. Once this was done, in theory schools had access to the YouTube for Schools global whitelist of 1800 "EDU Channels" as well as the option for admin-created playlists. *Id*. But in practice, the sign-up flow was broken and YouTube for Schools experienced "lots of outages." Ex. 1076 at 9257. (This was not surprising given that YouTube for Schools was a consistently underfunded project (referred to internally as a "perpetual 20% project"). Ex. 745 at 73.) As a result, YouTube for Schools was not widely adopted. Ex. 743 at 4725. Ex. 744 at 9512. In 2012, YouTube for Schools had only been adopted by 4% of approximately 14,000 school districts in the United States. *See* Ex. 750 at 5; Ex. 751 at 5. Further, in September of 2012, EDU-mode was rendered toothless when YouTube became accessible via the HTTPS protocol (https://www.youtube.com). This allowed students to easily sidestep EDU-mode's protective filter. *See* Ex. 743 at 4725; *see also* Ex. 750 at 5; Ex. 751 at 5.

In the summer of 2015, YouTube launched Managed Restricted Mode ("MRM"), advertising that

it would allow school administrators to select one of three access settings and, in doing so, limit the corpus of videos available to students accessing YouTube on their school network. Ex. 1078 at 85 ("admin sets MRM controls at the network-level (e.g. any computer on the school network is categorized as MRM"); Ex. 745 at 25 (depicting permission levels available to school implementing MRM). Google Classroom users also had the option to set access restrictions for managed Google accounts. Ex. 1078 at 85. But mere months after launch, YouTube internally acknowledged that MRM was "not meeting school's needs," Ex. 753 at 4419, because its "very basic filtering options" did not prevent students from receiving, uploading, or sharing non-educational videos or watching distracting ads. Ex. 753 at 4419; Ex. 740 at 132:2-133:2, 190:13-191:9. YouTube knew that, despite the importance of MRM filtering to school districts, its MRM safety features were "trivially easy" for students to bypass. Ex. 1011 at 1769.ECM (listing several ways students could avoid protections, such as [redacted]. In addition, the descriptions for the access settings (Strict Restricted, Moderate Restricted, and Unrestricted) were "too vague to be applied," causing frustration for administrators who were unable to explain to teachers and staff why certain videos were blocked. Ex. 1078 at 10. (Based on the date of this document in 2021, this problem remained an issue six years after MRM launched.) Finally, MRM did nothing to limit students' use of YouTube during the school day on personal smartphones over cellular networks.

YouTube also knew it was not a proven platform in the classroom context. Ex. 757 at 9082 ("The YouTube experience in K-12 schools is broken."); Ex. 740 at 60:23-61:1 (YouTube has no data showing that use of YouTube by K-12 students results in better grades); Ex. 758 at 36 (YouTube "isn't designed for formal learning."); Ex. 745 at 31 (admitting MRM was only meeting schools' "baseline need" and did not work with other tools like Classroom and Chromebooks); Ex. 740 at 262:1-263:12; Ex. 757 at 9014 (acknowledging MRM has been "historically under-resourced," resulting in "increasingly common user issues and emerging business risks.").

Indeed, YouTube recognized that its disorganized library of videos, along with the high potential for users to be distracted from intentional viewing of educational content, made YouTube problematic in the educational context and limited its ability to deliver structured learning. Ex. 740 at 169:15-170:19. A key problem is YouTube's "Watch Next" feature—which would routinely offer "recommendations [that]

don't align to a defined learning journey," making it hard for students to "stay on track." Ex. 755 at 7404. One internal presentation provided an example, showing a user watching a YouTube video about linear equations and then being recommended a "Will Ferrell hilarious acceptance speech." Ex. 745 at 74. As the presentation acknowledged: "Hard to do my homework when I could be watching Will Ferrell." *Id.* But YouTube also knew that allowing users to remain focused on their original learning intention meant fewer views for non-learning videos that YouTube might recommend—meaning less engagement and less Watch Time. Ex. 755 at 7404 ("If you remove non-learning things from the watch next of learning videos, then those videos will have far less reach.").

Despite knowing that its use case as an educational tool was unproven, and despite knowing its existing safety measures for schools and students were ineffective, YouTube embarked on a campaign to increase school-day and classroom usage *even more*. Specifically, YouTube set out to close what it perceived to be an 80 million hour difference in watch time between any given Thursday and any given Saturday. Ex. 753 at 4419. YouTube understood this to be an engagement frontier for the platform and a business "opportunity in education." Ex. 740 at 129:25–133:7-15. In furtherance of its goals, YouTube pursued partnerships with organizations like the National Parent-Teacher Associations (PTA) and Common Sense Media due to their reputation as highly credible sources of information for educators and parents. *See* Ex. 1002 at 250:5-251:5, *id*. at 71:8-11. YouTube then used those partnerships to distribute information to schools and normalize itself as part of the classroom experience.

YouTube's agreement with the PTA was a paid partnership. Ex. 760 at 6433 (in 2022-23, [redacted] for basic renewal plus [redacted]+ to add PTA grants and conference fees). The partnership provided YouTube with a direct line of communication to PTA members. In 2022, for instance, YouTube employees presented on their "efforts for younger kids, tweens, teens and parents" at an hour-long session at the annual PTA convention. Ex. 762 at 3364. YouTube also reached PTA members across the nation through email lists and editorials in the PTA newsletter, such as one about YouTube's Youth Principles drafted by James Beser, Senior Director of Product Management for Kids & Youth. Ex. 761 at 0318. While these communications described the dangers of various third-party content (for instance, "viral hoaxes") they conspicuously excluded information about potential harms presented by YouTube itself. *See* Ex. 1002 at 100:2-24, 111:7-25; Ex. 740 at 349:25-350:18.

Like the PTA, Common Sense Media is a non-profit organization that is well-regarded by parents and educators. Ex. 1014 at 7, 20. Its entire focus is children's media, Ex. 1002 at 261:17-20, and it serves as a clearinghouse for safe technology and media guidance for kids, Ex. 1014 at 17-19. Among other functions, Common Sense Media releases movie and TV ratings to assess what content is kid-appropriate, *id.* at 263:18-264:8, and provides digital literacy curriculums for K-12 students. Ex. 1079 at 2.

The high-level goal of YouTube's partnership with Common Sense Media (CSM) was to drive improved trust among parents and key opinion formers ("KOFs"), Ex. 1080 at cell C13—including regulators, Ex. 1001 at 183:22-184:2, and legislators, Ex. 1024 at 227:9-17. YouTube recognized that consumers are more persuaded by authentic, unsponsored feedback—an area where CSM excelled. *See* Ex. 1002 at 254:16-255:7. YouTube also understood that parents in particular react well to seeing CSM's green check mark and branding, which feels vetted and safe. Ex. 1014 at 7, 20. By partnering with CSM, YouTube positioned itself as a platform safe for use by children, including children under the age of 13. Among other promotions, the partnership resulted in two youth content creation guides (one for teens and one for parents) and sponsored content on YouTube-branded websites directed toward parents. Ex. 1080 at C14. Notably, despite a stated objective of "educat[ing] stakeholders and users," *id.*, none of this material refers to the risks of addiction and compulsive use that YouTube knew are inherent to its platform.

In December 2022, CSM and Frances Haugen co-authored an amicus brief filed in *Gonzalez v. Google* which, among other things, argued that "Section 230 does not grant blanket immunity" for a variety of Google's "non-publishing activities." Brief for Common Sense Media and Frances Haugen as Amici Curiae Supporting Petitioners at 3, *Gonzalez, et al., v. Google LLC,* No. 21-1333 (Dec. 6, 2022). Months later, CSM published a report on Teens and Pornography that was highlighted by the *New York Times*. Ex. 886 at 20, 22; Ex. 888 at 1. In response to these developments, Google prepared a dossier regarding its numerous business and financial entanglements with CSM and asked for a meeting to discuss their deals—including a [redacted] partnership with CSM's for-profit subsidiary, Common Sense Networks, and a Google TV partnership which cost Google [redacted] annually. Ex. 1111 at 0511-12 (talking points: "In June 2022 CSM did not hesitate to call Google out publicly, despite our active partnership deal."). YouTube's message to CSM was not particularly veiled: "We've heard feedback from

the teams that while they are not against working with Common Sense Media, there is hesitation around the scale, fit, and cost of the products." *Id.* at 0512.

Ultimately, YouTube's brand cleansing worked, and YouTube's incursion into schools was a success. Today, YouTube's CEO, Neal Mohan, is confident that more than 90 percent of teachers use YouTube in the classroom every single day. Ex. 1012 at 168:21-23. And still, YouTube has no data showing that the use of YouTube by K-12 students improves reading skills or any other skill, as YouTube does not measure the effectiveness of its efforts with respect to the use of YouTube in classrooms, Ex. 740 at 61:16-20; 64:24-64:6. In fact, what YouTube measures is users "really valu[ing] YouTube video"—YouTube's euphemism for engagement with the platform. *Id.* at 64:7-65:18; *see* Ex. 772 at 7889 ("satisfaction" still measured in part by Watch Time).

In 2022, after more than a decade of pushing its platform into the classroom space, YouTube launched a product called Player for Education. Ex. 740 at 43:17-44:5, 51:14-52:4, 162:20-163:15; *see also* Ex. 749 at 2037 ("full rollout" of Player for Education in mid-2023); Ex. 746 at 7. Player for Education enables teachers to assign specific YouTube videos to students via an embedded player within Google Classroom, or other Google products, with no links back to YouTube, no recommendations, no ads, and no cookies.[7] *See* Ex. 740 at 43:17-44:5. This solution attempted to address several problems reported by educators, while removing impediments to greater adoption of YouTube in classrooms. However, Player for Education does nothing to address the issue of student access to YouTube during school hours or on personal devices.

Finally, YouTube delayed implementing any screen time management controls effective during the school day until late 2024 (well after these lawsuits were initiated). Before then, granular school-day controls were not available, and parents had to attempt to devise workaround methods to limit use of YouTube during school hours. For example, in March 2021, YouTube employees noted in an internal

[7] YouTube's 30(b)(6) corporate representative on school district issues, Katie Kurtz, briefly mentioned an embed player pre-dating Player for Education. *See* Ex. 740 at 52:5-13. YouTube's iteration of the embed player that preceded the Player for Education was never a real solution for teachers because it was hard to use and, even when used, still showed "related videos." *See, e.g.*, Ex. 931 at 9197 (internal chat: "teacher must embed in a specific way" that even led Google staff to get "An error occurred' message); Ex. 1174 at 3620 (embed player used in Google Classroom played "related videos").

discussion that they were using the bedtime settings to block their children's YouTube access during the school day. *See* Ex. 766 at 476:21-477:8; 477:15-480:3; Ex. 764. YouTube did not launch a school time feature, to assist parents with controlling their children's YouTube access during school hours, until August 2024. *See* Ex. 766 at 470:2-20; Ex. 1002 at 205:15-206:10.

(4) YouTube's knowledge of harms

YouTube acknowledges that "kids [and] teens" are "vulnerable groups." Ex. 736 at 33. It knows that "brain development predisposes young teens to act more impulsively [and] show great tendency towards risk taking" and that "[e]xecutive functions aren't fully developed at this age (e.g., self-regulation and decision-making capabilities)." Ex. 735 at 5; *see* Ex. 1010 at 156:3-16; Ex. 737 (YouTube CEO: "[t]eenage brains are still developing and work differently than adult brains, particularly when it comes to decision-making or self-regulating"); Ex. 738 at 7 (identifying vulnerabilities specific to teens); Ex. 739 at 38 (citing a consistent increase of self-harm and suicide rates among adolescents and noting "Teens are more vulnerable because their judgment and decision-making abilities are still being developed.").

Despite this knowledge, YouTube designed its platform in a way it knows impacts youth wellbeing and promotes addictive and compulsive use. A 2016 presentation to executives admits as much: the YouTube Main App team "aspire[d] to create an app that is…Addictive[:] Our app experience should compel users to come back more and more often." Ex. 765 at 11. (In this same presentation, YouTube reported that it was reaching over six billion daily active views in the U.S. alone. *See id*. at 14.) Other presentations are similarly candid. *See also* Ex. 701 at 1255 (goal to "[i]ncrease habitual users" and "focus on making YouTube a daily habit"); *id*. at 1235 (discussing "aim of increasing frequency and long-term YouTube time"); Ex. 767 at 5098-5099 ("Watch time, and only watch time;" "All other things being equal, our goal is to increase (video) watch time"); Ex. 768 at 6246 (identifying notification features to "help build a YouTube habit[]"); Ex. 769 at 2 (presentation by YouTube Growth Team exploring "Principles of Habit Building" and brainstorming "features and concepts that build habits on YouTube").

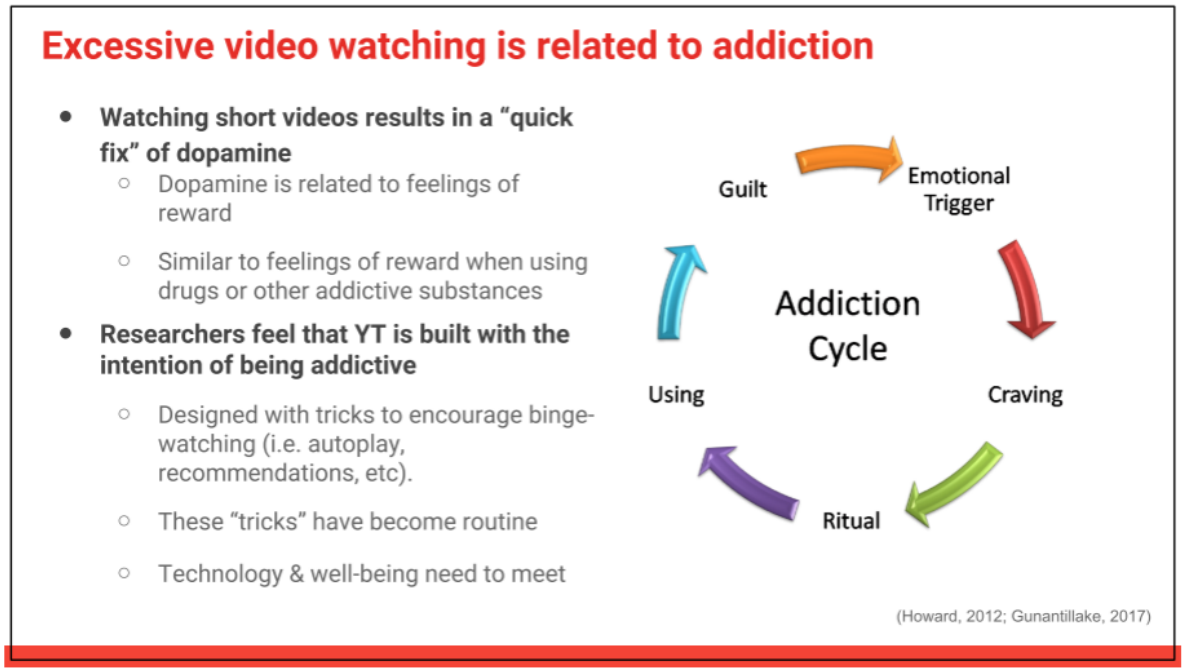
By at least 2018, YouTube knew that "tech addiction" had become a "major topic of concern," particularly the risks it presented "to the physical, emotional, and social development of teens." Ex. 772 at 7914. It identified three areas of concern impacting users 13-24 disproportionately:

- *Habitual heavy use:* ~10% (32MM) of 13-24 year olds on YouTube habitually[7] watch > 2 hours / day (excluding music).[8] ~13% (36MM) of 18-24 year olds reported "I regret how long I stayed on YouTube" in the past week.
- *Late night use:* ~7% of teens on YT watch past midnight on school nights. Teen "night owls" were 88% more likely to have emotional & behavioral problems (article, study). 30% of users 18-24 say YouTube has cut into sleep.
- *Unintentional use:* Among users 18-24 years old, 23% report "losing track of time on YouTube,"[9] 20% report "procrastinating on YouTube," and 20% report YouTube "interfered with work, school, or homework."

Id. Despite gaining an awareness of widespread "habitual heavy use," "unintentional use," and sleep-disruptive "late night use" by teens, *id.*, YouTube did nothing to warn the public or the Plaintiffs. Instead—within the very same document—YouTube focused on Meta's ability to "scale up monetization faster than YouTube." *Id*. at 7898. YouTube proposed solving this "problem," first, by continuing to "optimize for getting people to initiate long video watching sessions"; second, by introducing new feeds to promote "more scrolling," *id*.; and third, by developing a video format "optimized for the phone." *Id*. at 7960.

In 2018, YouTube conducted various studies and literature reviews that shed further light on the scope of the addiction problem. One of these reviews noted: "Research suggests that a common side-effect of using platforms with on-demand video can result in binge-watching and procrastination (Pena, 2015)." Ex. 773 at 19. "It is also known that YouTube's algorithms are designed in a way to increase 'rabbit hole' watch time and 'keep people hooked on the screen' (Lewis, 2018)." *Id*. at 19 (emphasis removed). Another defined "Technology Addiction" as the "widespread obsession, compulsive and excessive use of technology that interferes with Daily Life"; acknowledged research "indicat[ing] that [the] prevalence of tech addiction is inversely related to quality of life"; and observed that "Tech Addiction and Google's role has been making the news." Ex. 774 at 3. This same review flagged YouTube's autoplay feature as "Compulsive Content Consumption" and labeled it a disrupter of sleep patterns. *Id.* at 6, 8.

Still another literature review acknowledged that "excessive video watching is related to addiction"—and concluded that watching short-form videos gives users a "'quick fix' of dopamine" that is "[s]imilar to feelings of reward when using drugs or other similar substances."



Ex. 775 at 13; *see also* Ex.1081 at 290 (acknowledging that short-form content "exacerbates concerns of addiction and ADHD"); Ex. 1082 at 5298. (Trust & Safety employee asking whether YouTube has "any responsibility to moderate binge watching"). Despite its understanding that short-form videos can trigger an "addiction cycle," Ex. 775 at 4191, YouTube proceeded to develop Shorts (its short-form video format), which it released in the United States approximately three years later. Ex. 1112 at 6.

YouTube's negative effects on millions of young people—and the panoply of harms it could have and should have warned the public about—are laid bare in a presentation titled *Double Rainbow: A Vision of Digital Wellbeing at YouTube*. Ex. 738 at 2. In this document, which employees authored in 2018 but iterated on for years, YouTube acknowledged its negative effects on wellbeing, including "Sleep quality/duration," "FOMO," "Isolation," "Depression," "Anxiety," and "Body/Esteem issues." *See id.* at 15-16, 28 (evidencing comment from 2018). It admitted: "YouTube is designed around increasing users' engagement, not maintaining user's intention." *Id.* at 73. It observed: "Experts compare YouTube to a slot machine, with variable rewards and infinite possibilities." *Id.* at 77. In another version of *Double Rainbow*, YouTube concluded, based on user survey evidence, that 10% of users aged 13-24 watched 2 hours a day habitually, Ex. 1028 at 41, and 5% watched 3 hours a day habitually. *Id.* One 16 year old reported, "It delayed my sleep a lot. I'd watch until 2:00 or 3:00 in the morning...I'd be so invested that I'd keep watching." *Id.* at 9.

This same document estimated that 69 million teens were on YouTube. *See id.* at 41.

(5) YouTube's failure to exercise reasonable care

YouTube had an obligation to ensure that the YouTube Platform was appropriate and safe for young users—a demographic it was intentionally targeting. *See, e.g.,* Ex. 1013 at 27:3-7 ("if we learn about something that has kind of a cause-and-effect impact to youth, we would take action to…adjust our product experiences to make sure that it is most appropriate for younger users"); Ex. 714 at 45:16-46:3, 53:23-55:22 (YouTube's former Senior Director of Trust and Safety agreeing that companies should assess products for risks to users, that this obligation is ongoing, and that YouTube has a responsibility to ensure that its products are safe for children, and to warn parents and the public of any risks). In 2024, YouTube even codified some of these principles in writing, authoring a "Youth Legislative Framework" that would require online services to prioritize the best interest of children and teens in the design of their products. Ex. 770 at 2; *see generally* Ex. 771. YouTube failed to live up to these principles or obligations. To the contrary, it knowingly embedded a bevy of harmful, addictive features in the YouTube platform.

(a) Autoplay

Autoplay was launched on YouTube in 2015. *See* Ex. 789 at 1950. It is designed to create a seamless viewing experience, by automatically and continuously playing one video after another without any affirmative user action. *See id*. At the time of its launch, and up until 2021, the default autoplay setting for all users, regardless of age, was "on." *See id.* at 1952. This was a business decision, made without any concern for user wellbeing: "**We expected Autoplay on by default to generate more WT** [Watch Time]." *Id*. (emphasis in original); *see also* Ex. 790 at 9 ("More autoplay = more WT, and that was an explicit goal of the project"). YouTube was right. Shortly after launch on desktop (2015), autoplay was credited with generating 16% of YouTube's desktop watch time. *See* Ex. 791 at 3300. Its growth contribution continued through 2019, when it was credited with driving [redacted] of total Watch Time across all interfaces. *See* Ex. 789 at 1954. YouTube has internally lauded Autoplay as "the single most impactful launch in YouTube history." *See* Ex. 765 at 28.

YouTube knew that autoplay contributed to excessive and addictive use. *See, e.g.*, Ex. 789 at 1962 (acknowledging autoplay creates wellbeing issues for some users, which includes prolonged unintentional

use and sleep disruptions); Ex. 790 at 10 ("Autoplay's share of watch time doubles at night"); *see id.* at 12 (acknowledging concerns about autoplay from media, key opinion formers, and legislators); Ex. 775 at 13 ("Researchers feel that YT is built with the intention of being addictive [and] [d]esigned with tricks to encourage binge-watching (i.e. autoplay…)"); *see also* Ex. 987 ¶¶ 59-61, 200-202; Ex. 982 ¶ 193; Ex. 1004 ¶ 445; Ex. 983 ¶ 5.b. Nonetheless, it failed to inform its users, their caretakers, Plaintiffs, or the public at large that autoplay on YouTube contributes to excessive and addictive use.

In 2021, YouTube turned autoplay *off* by default for "all users 13-17." Ex. 1064 at 0252. It also added an option to toggle autoplay on or off on YouTube Kids, setting the default *off*. *Id.* This was an important measure. But it came with some serious limitations. First, the default setting only applies to users who are logged-in; and, as discussed below, YouTube is easily and often accessed in a logged-out state. Second, "logged-in" teenagers will only have autoplay defaulted to *off* if they correctly identified themselves as teenagers upon sign-up. YouTube knows this does not happen for millions of teenage users and (again, as discussed below) it intentionally delayed implementing age inference models to gain a robust understanding of user age. Ex. 1028 at 41 (YouTube estimates there are 69M teens on YouTube); Ex. 1137 at 78 (YouTube asserts that actual teens on the platform are close to [redacted] of daily average viewers rather than the 4% who declare); Ex. 719 at 3742 (only a small fraction of those who are actually under 18 are accurately declaring their age); Ex. 932 at 0478 (2022 email chain discussing [redacted] and noting that "there are still [redacted] underage accounts left in the corpus at [redacted] precision"). Third, YouTube employed dark patterns to encourage users to turn autoplay back *on*, designing the toggle to be prominently displayed and easy to access. *See* Ex. 1010 at 288:25-289:11, *id*. at 292:2-12 (autoplay toggle was "very prominent" and "easy-to-understand"; it was "highly visible…right on the player in mobile" and "right below the player" in desktop). By contrast, and as discussed in later sections, *see infra* § III.A.2.d.h, YouTube's screen time management tools are buried in a settings page.

(b) Infinite Scroll

YouTube employs endless scroll features on its platform, including on homepage, Watch Next, and Shorts. *See* Ex. 792 at 5068; Ex. 793 at 6563; Ex. 1010 at 230:25-231:9, *id*. at 233:16-235:1; *see also* Ex. 987 ¶ 203. For example, YouTube Shorts is designed to display in a "Shorts Player," an infinite feed of content modeled after TikTok and Instagram Reels. *See* Ex. 1172 at 1649, 1665 (internal TikTok

document from 2020 describing "Short-video war" and noting that "users can swipe through YouTube Shorts vertically—just like TikTok."). YouTube is aware that endless scroll features contribute to compulsive and addictive use. *See* Ex. 715 at 351:9-23 (moving Watch Next feed to infinite scroll increased user time on platform); Ex. 796 at 8133 (experts highlighted that endless scrolling can be detrimental); Ex. 732 at 9252 (concerns about prolonged unintentional use "are loudest on short-form content (more popular with teens) due to its lack of depth and infinite feed experience").

This is echoed by the testimony from Plaintiffs' experts. *See* Ex. 987 ¶¶ 59-61 (identifying features like infinite scroll as a type of dark pattern designed to maximize engagement and attention); Ex. 1004 ¶ 445 ("YouTube's design features, such as likes, comments, notifications, autoplay, and endless scrolling, encourage addictive screen time."); Ex. 983 ¶ 5.a ("continuous scroll takes advantage of multiple aspects of brain development, less developed self-regulation, adolescents' focus on social connection and peer feedback due to social development, and adolescent egocentrism"); Ex. 982 ¶ 194 ("Endless Scroll feature has been linked to prolonging engagement at the cost of displacing children's developmental opportunities").

Finally, as discussed above, the inventor of the infinite scroll, Aza Raskin, has publicly stated that the goal of the feature was to "deliberately keep [users] online for as long as possible." Ex. 252 at 2. At his deposition, he described the infinite scroll as "an intentional removal of stopping cues so that your brain doesn't wake up to catch up with impulses….it creates that hypnologic state where you just keep scrolling." Ex. 1070 at 41:16-21; *see id.* at 55:21-56:5 (reviewing example of infinite scroll on YouTube). He testified that, "as the designer of infinite scroll, I know that if you do not give the user's brain a chance to catch up with … their impulse, then you can predictively get them to continue to scroll and look at content." *Id.* at 57:18-58:6; *see id.* at 55:21-56:5 (reviewing example of infinite scroll on YouTube). "If you had a glass of wine that just kept refilling, like, your mind would never have that moment where you say, oh, I reached the bottom, do I really want another or are you just going to keep sipping and keep sipping." *Id.* at 24:22-25:1. Sites like YouTube "are all taking great care to preload the content so that the infinite scroll is as smooth as possible." *Id.* Reflecting on how his invention has been deployed by social media platforms, Raskin stated: "when you get on to a bus or a train or you go into a classroom …and you watch their thumb flicking and flicking and flicking…because of…an invention that I had made that is

now being used for something else, it's very hard. Like that is a deep, hard, sick place in the pit of my stomach when I really get in touch with it." *Id.* at 65:6-18.

YouTube has failed to present any evidence in its summary judgment brief that, at any point, it informed its users, their caretakers, Plaintiffs, or the public at large that endless scroll on YouTube contributes to excessive and addictive use.

(c) Algorithm

The YouTube recommendation algorithm was launched in 2008 and determines what content will be recommended to users on YouTube. Ex. 1045 at 9. It is designed to keep people on the platform by presenting them with content that it infers (based on an analysis of behavioral signals) will keep them using. Ex. 546 at 4852 ("The user behavior patterns that are of our interest here are the ones we can *shape* through changes in the recommendation algorithms." (emphasis in original)). Ex. 1045 at 11 (recommendations are made by leveraging user-specific and video-specific signals, including the user's watch history, search history, demographics, and activity level); Ex. 776 at 6773 ("There is a lot more we can infer from user's interactions—e.g. what items they scroll past without clicking—and update results in real time.").

For many years, YouTube explicitly optimized its recommendation systems to maximize Watch Time above all else, even when doing so meant "essentially ignoring the initial intention the user had when coming to YouTube." Ex. 767 at 5100; *see also* Ex. 1045 at 9-13; Ex. 767 at 5098-5100; Ex. 711 at 3; Ex. 985 ¶ 125; Ex. 767 at 5100 ("All other things being equal, our goal is to increase (video) watch time."). To satisfy YouTube's hunger for its users' time, the algorithm must be constantly updated—hundreds of small changes that add up to large effects on users. YouTube uses an experimentation platform called [redacted] to measure the effectiveness of these changes. For instance, in 2018, there were [redacted] changes made to the recommendation system, resulting in a [redacted] increase in "YouTube Time" (the largest component of which is Watch Time). Ex. 539 at 46. YouTube's users are also its lab experiments.

Since 2016, YouTube has included user "satisfaction" alongside watch time as a metric that its recommendation system optimizes for. Ex. 985 ¶ 132. However, as the Districts' algorithm expert makes clear, YouTube measures "satisfaction" in ways that often bottom out in engagement metrics like time spent. *See id.* ¶¶ 132-137. Similarly, certain YouTube witnesses testified about the company's attempt to

"drive value" for users, including through its recommendation system. Ex. 1003 at 41:2-10, 326:08-14, 334:10-23. However, YouTube's employees understood that "the stated goal of value was viewer addiction." Ex. 1116 at 0356 (noting this in section titled "Value prop"); Ex. 1003 at 195:24-196:03.

In short, regardless of the words used to describe it, the goal of YouTube's recommendation algorithm has been and continues to be keeping users engaged, i.e. watching videos. *See, e.g.,* Ex. 1045 9, 13-14; Ex. 772 at 7889 ("engagement" is still measured primarily by Watch Time or YouTube Time and focuses on "driving greater consumption of videos"); Ex. 982 ¶ 189 ("algorithms are optimized to maximize time on the platform rather than healthy interactions with a person's social network").

Recommendations surfaced by YouTube's algorithm are placed in areas that will draw the most user attention—Home, Watch Next, Search, the Shorts player, and YT Kids Home. Ex. 985 ¶ 123. They are displayed via autoplay and endless scroll features that drive over-consumption. *See supra*; *see, e.g.*, Ex. 775 at 3; Ex. 788 at 15, 33 (Watch Next panel "encourage[s] viewers to watch more videos" and "has been referred to as a 'rabbit hole'"); *see also* Ex. 983 ¶ 5.d. ("Recommender algorithms take advantage of multiple aspects of brain development as well as adolescents' less developed self-regulation").

YouTube failed to assess its recommendation system for risks to teens. It was not until 2021 (thirteen years *after* launching the system), that YouTube consulted with experts to identify potential mental health risks to teens. In doing so, YouTube learned that its algorithm can negatively impact teen wellbeing by aggregating and repeating certain types of videos, any one of which could be harmless on an individual basis but harmful in volume. Ex. 1026 at 2617-2620 (reiterating that the problem is watching volume of videos that repeat and reinforce skewed body norms; "too much of this content may lead teens to normalize this and feel pressured to do the same thing because their perception of reality is distorted").

Critically, the root of the harm stems not from the content displayed but from the YouTube recommendation system itself—which, like the recommender systems of YouTube's co-defendants, has a proclivity towards "strong feedback loops." Ex. 1173 at 31; *see* Ex. 985 ¶ 183. This means that, "[o]nce you watch a few" of a certain kind of video, "your feed might become concentrated [with] a high volume of content that repeat the same message." Ex. 1026 at 2619-20. YouTube understood, for example, that social comparison and body image videos, when viewed in repetition/volume, can skew teen perception of norms and lead to harmful behaviors. *See* Ex. 1026 at 2617-2620 ("sometimes the Watch Next feed

may have many more videos that repeatedly make negative social comparisons of physical features."); Ex. 1025 at 2611; Ex. 1028 at 2014, 2017; *see also* Ex. 1023 at 95:21-25 (former senior product manager for health and mental health: "I would say it's a combination of the user kind of urged to watch or choice [sic] to watch more content this way. And then the loop of it feeding that signal into the recommendation system"). In other words, the algorithm was tuned to do whatever it took to keep children's eyes on YouTube, without any regard to what was being shown. But the inevitable outcome of this behavior-maximization was children becoming trapped in echo chambers of ever escalating severity.

For example, YouTube understood that a person watching a video presenting a high degree of social comparison would have a 45% chance of receiving Watch Next recommendations for two similar videos. Ex. 1025 at 15. Experts advised that having just three or more such videos recommended in a single session is "too many." *Id.* at 16. Internally, YouTube acknowledged that "[t]hese concerns have also been escalated in the news by parents and [key opinion leaders]." Ex. 1026 at 2625.

YouTube's recognition of the harms caused by its algorithm is further evidenced by Project VIBE ("Volume Impacts Well Being"), whose goal was to ameliorate the harms caused by this algorithmic clustering. Presented on August 10, 2022 by the Youth, Health, and Trust and Safety teams, the initiative had two parts: first, create a classifier to identify categories of content recognized as harming the wellbeing of teens when viewed at high, repetitive volumes; and second, use that classifier to disaggregate harmful algorithmic recommendations and limit its overwhelming rabbit hole effects. Ex. 1028 at 3, 18, 64, 68, 76, 117-118; Ex. 1025 at slide 19; Ex. 1001 at 74:4-I10; Ex. 702 at 45:6-50:25. Discussing the effects of VIBE, Dr. Jessica DiVento Dzuban, formerly YouTube's Global Head of Mental Health, offered the "[h]ypothesis …that exposure without VIBE would increase risk. But doing A/B testing on this is likely unethical (as we wouldn't expose someone to risk increasing content if that's our hypothesis)." Ex. 1024 at 25:10-13; Ex. 1029 at 212:22-213:2. Stripped of research jargon, this is a YouTube executive hypothesizing that the experience of YouTube without Project VIBE—i.e. the experience that children had received for over a decade—put children's wellbeing at *so much risk* that that it would be unethical to continue exposing children to it in order to formally test the hypothesis. Indeed, there is no evidence that A/B testing was ever used to measure the efficacy of VIBE.

While the existence of Project VIBE demonstrates that YouTube was aware of the risks its algorithms posed to young users, its actual implementation is evidence of YouTube's broader failure to take these risks seriously. When it launched, Project VIBE was initially limited to the Watch Next panels. It did not launch on Shorts—YouTube's most popular surface for teens—until Q4 of 2023. *See* Ex. 702 at 45:6-50:25; Ex. 1029 at 234:1-24. Even then, it only applied to Declared Teens who were *also* logged into their account at time of use. *Id.*; *see also* Ex. 545 at 16 (stating "[d]eclared Teen has low recall. Only [redacted] of teens on [YouTube] declare 13-17" and "[d]eclared teens only ~3% of [Daily Active Users]").

YouTube did not inform its users, their caretakers, Plaintiffs, or the public at large that its recommendation system was aggregating content in harmful ways and at harmful volumes—nor that its attempted solutions were limited in scope and did not extend to Shorts. *See* Ex. 1023 at 156:10-19.

(d) Age Verification

The Children's Online Privacy Protection Act (COPPA) requires platforms like YouTube to, among other provisions, obtain verifiable parental consent if they allow users under the age of 13 on their site. 15 U.S.C. § 6502(b)(1)(A)(ii). If they allow such users, they must also provide parents with the opportunity to review personal information collected from their children and the opportunity to refuse further use or maintenance of such information. 15 U.S.C. § 6502(b)(1)(B). These are in addition to a bevy of other requirements. *See, e.g.*, 16 C.F.R. § 312.4 (requiring posted privacy policies); 16 C.F.R. § 312.5 (mandating mechanisms of parental consent); 16 C.F.R. § 312.6 (establishing parents' right to review information provided by a child); 16 C.F.R. § 312.8 (imposing data retention and deletion requirements). In its summary judgment briefing, YouTube presents no evidence that it complies with these requirements. If anything, YouTube appears to rely solely on the fiction that it does not allow children 13 or younger on its platform. *See* Exs. 1047, 1067, 1118, 1119, 1120, 1121 (YouTube Terms of Service from 2018 through 2023); Ex. 745 at 32.

In reality, a key element of YouTube's mission to embed itself in the daily lives of young users is the ease with which children can access YouTube. Anyone—regardless of age—can access YouTube. No account, screening, age information, or guardrail of any kind is deployed by YouTube to prevent general access. And YouTube is well aware that this unfettered access leads to large numbers of children under the age of 13 (u13) using the platform. *See* Ex. 724 at 5688 ("vast majority of u13 usage is…happening

anonymously on [YouTube]," e.g., without an account); Ex. 766 at 491:23-492:1 ("YouTube signed out is very easy to access, and it's very likely that some of those users are under 13"). YouTube's failure to effectively prevent children under the age of 13 from accessing the platform, at any time of day, has compounded harm to schools and school districts. Ex. 1022 at 86:19-22, 251:4-20 (noting YouTube was aware it was a source of distraction for students).

YouTube makes certain features such as commenting, liking, and subscribing available only to users who are logged in. Ex. 716 at 59:21-60:16 (describing pop-out screen prompting sign-in). But YouTube entices kids to do precisely that, by ***actively prompting*** children to create an account and failing to implement any meaningful age gate during the account creation process. *See* Ex. 987 ¶¶ 42-62; Ex. 715 at 324:3-21 (testifying that the logged-in experience had to be more compelling to discourage logged-out usage). Prior to 2016, there was no age-barrier to account creation. Ex. 717 at 1, 6, 8. Since 2016, YouTube has required that users input a date of birth of 13 or older (the "Declared Age") to create an account. *Id.*; Ex. 718 at 7653. But there is no age validation measure to confirm the birthdate provided, meaning children can simply input an inaccurate birthdate and immediately unlock the full YouTube experience.

YouTube is well aware that Declared Age is unreliable. *See* Ex. 719 at 0742 ("only a small fraction of those who are actually u18 are declaring accurately"); Ex. 720 at 3366 ("Declared isn't a reliable signal to build teen specific models); Ex. 722 at 8465 ("We've heard legitimate concerns that kids under 13 are using the main YouTube site unsupervised.…**We know it is our responsibility to address how our products are being used**" (emphasis in original)). But YouTube failed to implement age validation, knowing it would conflict with its goal of obtaining and retaining young users. *See, e.g.*, Ex. 721 at 5186.ECM (acknowledging that "collecting a date of birth (DOB) at sign up for an account is no longer sufficient for products that are deemed not suitable for children" and "Google has to come up with a comprehensive strategy to tackle this problem ***while ensuring we do not negatively impact the user experience or our products***") (emphasis added).

YouTube's age verification failures cannot be ascribed to technical capacity, and instead reflect the systemic, years-long deprioritization of safety, even in the face of internal alarm. Since at least 2018, Google's advertising teams have been able to infer age for advertising purposes. *See* Ex. 1042 at 48:7-14, 59:18-20; Ex. 909 at 34:6-35:8. This is because advertisers want to understand who they can target, and

certain products are more suited towards certain people, certain genders, certain ages. *See, e.g.,* Ex. 1010 at 50:19-51:10 (interest-based marketing is where "we build an understanding of what interests a user might have based on…where we have visibility into them across the web, and we create interest categories that we associate with those users"); Ex. 909 at 35:17-36:9 (inferred age in Google Ads available "for quite a while now"); *id.* 64:25-65:3, 66:8-16; Ex. 781 at 2448 (document referencing data column [redacted] which is age inference). Yet prior to 2019, YouTube's only method of detecting and removing underage users was to manually remove them when employees, through pure happenstance, became aware that a user was under 13 in the course of carrying out other work. *See* Ex. 1032 at 2. Automated processes could be used to sell things to users, but not to keep them safe.

In 2019 YouTube finally created a machine learning classifier to detect u13 accounts, which it called [redacted]. *See id*. 2. In theory, once [redacted] flagged an account as u13, the account would be "auto-actioned." But YouTube placed a number of restrictions on this process, significantly undercutting its efficacy. ***First,*** even if accounts are flagged by [redacted], the "auto-action" is not deletion. Instead, YouTube sends the users an automatic notice informing them they have 14 days to confirm they are either under 13 years of age (in which case, they must set up a supervised account) or over 13 years of age (by submitting proof of, for example, a government ID), or delete the account. *See* Ex. 780 at 2810; Ex. 718 at 4. ***Second***, [redacted] review was limited in scope. [redacted] From November 2019 to November 2020, the u13 classifier [redacted], and from Feb. 2021 to Jul. 2021 the classifier [redacted]. *See* Ex. 780 at 2809. This limitation meant that [redacted] ***Third***, YouTube did not include [redacted] in the scope of its classifier's review, despite knowing that a [redacted] *Id*. at 2813 (2023 document discussing need to include [redacted] in u13 classifier review). ***Fourth***, [redacted] *See* Ex. 1122 at 87:19-23; Ex. 1123 at 1. From November 2019 to July 2021, flagged accounts would only be [redacted]. Ex. 1032 at 4; *see also, e.g.* Ex. 780 at 2810; Ex. 718 at 4. In July 2021, that threshold was [redacted]. Ex. 1032 at 4. Anything lower—[redacted]

[redacted]

YouTube also placed a [redacted] Ex. 780 at 2812; *see also* Ex. 1032 at 4. YouTube knew that these [redacted], with internal reports warning that [redacted] Ex. 780 at 2812 (emphasis added). For example, in 2024, YouTube found that "underage accounts that are actioned by [redacted] are generally [redacted] Ex. 780 at 2812 (emphasis added). YouTube further recognized that [redacted] Ex. 780 at 2813. The [redacted] therefore reflects critical flaws in u13 review: young users were allowed to remain on the platform [redacted] and thus did not benefit from any age-verification measures to correct their age. But rather than invest resources in [redacted] Google—one of the wealthiest and most successful technology companies in the world—left the [redacted] in place. *See* Ex. 1032 at 5. YouTube also never informed parents, teachers, or schools of the fact that its age-verification process was intentionally designed to be attenuated and inefficient.

YouTube has the capacity (if not the will) for effective age verification. In 2020, the EU amended its Audio Visual Media Services Directive (AVMSD), requiring YouTube to "establish and operate age verification systems to protect minors or under 18 users from content which may impair the physical, mental or moral development." Ex. 782 at 8. This rendered inadequate YouTube's approach of allowing a user to input a Declared Age to access mature content. *See id.* at 11-12. So YouTube began developing an age inference model to detect under 18 year old users, which it launched in the EU in late 2020. *See* Ex. 782 at 12. Among other things, this system leveraged Google demographic signals. *See id.* at 12; Ex. 717 at 1.

In September 2021, the United Kingdom's Age Appropriate Design Code (AADC) came into effect, after a twelve month transition period. Ex. 1141 at 16. The AADC required YouTube to "provide either a child-appropriate experience for all users, or apply an 'age assurance' model with a sliding scale

of different levels of age verification depending on nature and risks associated with processing." Ex. 1142 at 14. Faced with a requirement in the UK to "verify the age of all users or apply requirements to all age unknown users," Ex. 783 at 8, YouTube chose the former option, and deployed an age inference model to detect users under 18. *See* Ex. 717 at 1. For users in the UK, as well as the European Economic Area and Switzerland, YouTube began to use inferred age to apply age-appropriate design features discussed elsewhere in this section (e.g. take a break reminder, bedtime reminder, the VIBE classifier, turning autoplay off by default). *See* Ex. 783 at 28-29; *see also* Ex. 880 at 9452 (geographic area for "Minor Mode"). This age assurance model did not require any new signals and data collection points. *See* Ex. 717 at 2; Ex. 716 at 32:23-33:2.

Meanwhile, back in the United States—where YouTube remained free and clear from such regulatory requirements—it deployed none of these *already available* age inference measures to protect American teenagers and children. Instead, YouTube continued to rely solely on Declared Age, which it knew to be unreliable. Ex. 719 at 3742. YouTube's decision to do nothing more for American users is egregious and can only be explained by the company's determination to prioritize revenue over user safety.

This only began to change in February 2025—well after the initiation of this lawsuit. Ex. 1012 at 192:20-194:21 (testimony from CEO Neal Mohan that YouTube was, at this point, "willing to take that financial hit" because "the models have gotten to the point where we could have a pretty good inference model."). On February 11, 2025, YouTube announced that it *planned* to "use machine learning in 2025 to help us estimate a user's age – distinguishing between younger viewers and adults – to help provide the best and most age appropriate experiences and protections." *See* Ex. 717 at 18. YouTube's incredibly belated announcement that it would apply age inference to U.S. users, s*ee* Ex. 716 at 53:9-54, came five years after it had already launched age inference models in the EU. And it came after *years* of internal discussion concerning the need for a global model. *See* Ex. 1044 at 186.ECM; Ex. 784 at 6550; Ex. 719 at 3742; Ex. 785 at 3214; Ex. 786 at 18.

YouTube has tried to excuse its delay on the grounds that age inference models are "technically challenging." Ex. 766 at 553:14-554:4. But, as discussed above, its parent company Google already had the data points needed to infer age, having done so for years to make additional profit in the advertising space. And YouTube had already launched age inference models in other jurisdictions for the express

purpose of inferring users under the age of 18. Technical feasibility is simply not an excuse.

YouTube has also maintained that, because age inference models are probabilistic, they needed to be tested and configured to ensure an appropriate level of accuracy. For example, YouTube's CEO expressed concerns that incorrectly inferring an adult user as a teen would lead to a degraded experience for the adult because, for example, they might not receive personalized advertising. *See* Ex. 1012 at 199:11-15 (attributing delay in part to concern that "some number of users who might be above 18 are now for some reason classified as below 18. And so, therefore, they get a degraded [non-personalized] ad experience"); *see also* Ex. 766 at 677:13-679:17 (concern from senior director that false positives would be disruptive and annoying for adults) .

But this parade of horribles is a red herring and, even if true, a self-constructed barrier to avoid implementing age inference in the United States. YouTube can (and *has*, in response to the AADC and AVMSD) applied the output of its age inference models for the limited purpose of turning on protective settings for young people—which is beneficial, not harmful, to the overall user experience. Moreover, inherent in this excuse is the implication that perfection should supersede protection. In YouTube's world, the risk that it could potentially over-designate adults as minors—and apply protective features to them that would ***moderate usage***—is simply too high a cost. This reasoning is in direct contradiction with what YouTube itself recognizes is its obligation to protect young users from harms on its platform.

The real reason for YouTube's delay is more straightforward: YouTube feared a loss of revenue. In the run-up to YouTube's February 2025 announcement, the company forecasted the potential revenue implications of rolling out age inference globally. It determined the cost could range from [redacted] depending on the amount of "readiness work" that preceded the change. Ex. 787 at 8_022; Ex. 1012 at 206:5-18.

(e) Barriers to Exit

Even as YouTube made access to its platform incredibly easy—including through a logged-out state that requires no age verification—it has made account *deletion* incredibly difficult. *See* Ex. 987 ¶¶ 219-222 (YouTube account deletion process exemplifies "high-level dark patterns, which impede users from exercising autonomy"). In fact, the process for deleting a YouTube account is so difficult that in one "Google Account Deletion Flow" video produced by YouTube in this litigation, a system engineer was

unable to successfully delete an account—instead creating, and then deleting, an additional Google+ account. *See* Ex. 723 (video). The marked difference between ease of access, and difficulty of deletion, demonstrates that YouTube can and does leverage design barriers to serve YouTube's bottom line.

(f) Parental Controls

As detailed in this section, YouTube's purported parental controls and supervision tools were, and remain, woefully deficient in terms of addressing addictive and compulsive use.

Family Link (launched to the public in 2017, *see* Ex. 1143 at 1942) is not a YouTube specific parental control/supervision feature. *See* Ex. 730 at 3; Ex. 766 at 472:8-22. It is a complicated set of controls that is limited to Google accounts only; that is buried in the Google accounts setting page; and that varies across user devices (Android vs. iOS) and platforms (mobile versus desktop). *See* Ex. 1015 at 1210; Ex. 730 at 3. This makes Family Link ineffective as a screentime monitoring or supervision tool for YouTube, a fact YouTube acknowledges. *See* Ex. 1016 at 4601 ("Google and YT's current offerings are inadequate"); Ex. 1015 at 1210 (depending on device, children "can only be partially supervised"). YouTube knew that usage of Family Link was low. Ex. 1017 at 9 ("Family Link among parents is low [redacted]. It knew that variations across devices meant gaps in service. Ex. 1016 at 4601 (Family Link controls do not apply to the [redacted] of U.S. users on iOS). And it knew that options buried in account settings were difficult to access. *Id.* Furthermore, YouTube understood that children could easily "circumvent policy restrictions and parental controls" of Family Link by simply accessing YouTube in a logged-out state (i.e. not signed into an account). Ex. 1018 at 0294.ECM. This is assuming teens even opt in to Family Link in the first place—as the setting requires their agreement. *See* Ex. 1019 at 126:8-25.

Importantly, Family Link did not provide parents with an adequate way to monitor and control their children's usage of YouTube; as such it does nothing to help parents address addictive and compulsive use. *See* Ex. 1016 at 4601.

Supervised Experience on YouTube Main (SupeX) was launched in 2021. *See* Ex. 730 at 3. As discussed above, SupeX was designed to address a perceived product gap for 8 to 12 year olds aging out of YouTube Kids and seeking an older experience. *See, e.g.*, Ex. 731 at 6489 ("SupeX has the content missing" from YouTube Kids). SupeX was and is ineffective as a supervision tool. As with Family Link, it is easily circumventable: Anyone can access YouTube without an account, and tweens can easily create

an account without parental supervision. *See supra* § III.A.2.d.5.d.

Moreover, the supervision aspects of SupeX are woefully deficient in terms of addressing addictive and compulsive use. SupeX controls are focused on content control—e.g., choosing a specific content "corpus," channel blocking, and disabling or deleting a child's search and watch history. *See* Ex. 730 at 3-4. Those are all content-focused restrictions; they do not assist a parent in navigating the content-agnostic problem of addiction. And while SupeX does allow parents to turn off autoplay, it retains many of the other addictive features described in this section, such as an algorithmic recommendation system, notifications, filters, and an infinite scroll in Shorts, Watch Next, and elsewhere. Moreover, while YouTube knew that parents wanted some ability to control their children's screen time, they did not include any screen time limit within SupeX. *See* Ex. 1007 at 6134-35; Ex. 1008 at 5271 (2022 report indicating that parents thought a "Supervised channel" was one that let them "limit screen-time use").

Parent Supervised Experience for Teens was not launched until 2024. *See* Ex. 730 at 4-5. It is ineffective as a supervision tool. *First*, it is an opt-in experience for the teen; they can stop it at any time. *See* Ex. 1020 at 9045; Ex. 1021 at 211. *Second*, even if opted-in for one account, teens can circumvent the supervision by just creating other accounts. *Third*, with respect to addressing addictive and compulsive use, Parent Supervised Experience for Teens does nothing. It has limited tools only for "teen users who post content." Ex. 730 at 4. In those cases, parents will be able to see channel activity and receive a notification when the teen posts public video or starts a livestream. *See id.*; Ex. 1019 at 31:20-32:1. There are no controls as to screen time or any visibility for parents into their teen's watch time history. *See* Ex. 1031 at 4; *Id*. at 28:4-30:14, 87:12-25.

Given that YouTube's parental controls don't actually give parents control, it's fair to ask what they're actually for. The answer? To fulfill YouTube's business objectives. As noted above, SupeX was designed principally to help YouTube capture the market of "tweens" too old for YouTube Kids. *See supra* § III.A.2.d.2. Likewise, Family Link helped Google acquire young users by creating a path for them to create a Google Account (otherwise not permitted for those under 13). *See* Ex. 1139 at 9-10, 25; Ex. 765 at 42 (Deep Dive with Neal and Matthew) (noting "parents can use Family Link to create accounts for kids under 13"). While presented under the guise of "protection," supervision features like these were in fact just tools for YouTube to profit more off kids.

Finally, a critical component of effective control and supervision is to inform the decisionmaker of the impacts of their choice. YouTube never warned parents or kids about the addictive designs on its platform, and it never warned parents of the alarming gaps in the controls and supervision tools presented. *See infra* § III.A.2.d.7. Individuals cannot judge the level of vigilance and action needed if they are not informed of the scope of harm and if they are misled about the efficacy of the protections offered.

(g) Notifications

Notifications are a "critical" part of YouTube and an important way it drives engagement. Ex. 775 at 14; Ex. 798 at 0466 (describing use of "reengagement emails targeted at less engaged users."). YouTube's commitment to reengaging users through notifications is tremendous. Ex. 1051 at 216:21-218:21 (increase in notifications a core driver and "unlock" for engagement growth and business value). In 2018, its Growth Team listed as a top objective growing notifications' reach and impact, setting a goal of increasing daily notification openers to [redacted] and daily active viewers from algorithmic notifications to [redacted]. *See* Ex. 1048 at Growth Team Tab, Row 4. By 2021, YouTube was sending [redacted] notifications daily—up from [redacted] notifications per day sent in 2016. *See* Ex. 1049 at 9. This amounted to [redacted] algorithmically-generated, growth-oriented notifications. *Id.* at 9.

YouTube is aware that notifications contribute to excessive and addictive use, as they prey on users' "fear of missing out." *See* Ex. 775 at 6 (wellness factors diagram showing a connection between addiction, notifications and FOMO); *id.* at 10, 14 (receiving notifications causes users "to feel like they must be aware of what is happening on the platform" and tempts users to watch videos the moment they are uploaded). Self-reports from female users confirmed that morning and evening app use was triggered or driven by notifications, especially for teens, and that, relatedly, bedtime triggered notification checks commonly resulted in "mindless scrolling…before falling asleep." Ex. 1050 at 4675-76; *see also* Ex. 799 at 5968 ("Notifications can be distracting and overwhelming, especially close to bedtime"). Notably, despite YouTube researchers finding that "notifications... contribute to addiction," YouTube sent notifications to students during instructional school time. Ex. 775 at 10, 14. YouTube also never warned the public of the risk that its notifications posed to young users' wellbeing and instruction. Ex. 997 ¶¶ 155, 159-160, 166-67, 186-187, 191-92, 195-196, 201, 202.

The Districts' experts agree that YouTube's notifications present significant risk to teenage users in particular. *See* Ex. 1004 ¶ 52 ("Notifications and alerts that prompt repeated engagement leverage the variable timing of cues and rewards to reinforce compulsive checking behaviors"); Ex. 987 ¶ 53 (notifications are a type of dark pattern and "attention capture damaging pattern" that cause reactivation of engagement); Ex. 983 ¶¶ 5.c ("Push notifications take advantage of multiple aspects of adolescent brain development, less developed self-regulation, social development, and adolescent egocentrism").

(h) Screen Time Management

In 2018, YouTube launched a set of three tools to address digital wellbeing: (1) Take a Break Reminder, (2) Time Watched Profile, and (3) Notification Clusters. *See* Ex. 1033 at 0058. In 2020, it launched another similar tool, (4) Bedtime Reminders. *See* Ex. 730 at 5; Ex. 1034 at 9501. All four are opt-in tools, meaning that to access the tools, users would first have to know of their existence, then navigate to a separate setting screen, find the setting, and opt-in to receive the treatment. This high user friction ensured that few would opt-in, deeply limiting the tools reach and efficacy. *Cf.* Ex. 989 ¶ 34; *see also* Ex. 1022 at 199:7-201:25; *see generally* Ex. 1036; Ex. 987 ¶ 205.

Take a Break Reminder. Take a Break Reminder was aimed at heavy habitual use. *See* Ex. 1034 at 9499. At launch, it was an opt-in setting that allowed users to set a reminder to take a break after *n* minutes of watching videos. *See* Ex. 799 at 5966; Ex. 1034 at 9499. In addition to being hard to find, even if enabled it was easy to dismiss. Indeed, "the reminders are presented in way that *incentivizes* dismissal and continued content consumption, rather than choosing to end a session." Ex. 987 ¶ 206 (emphasis added). That is so because, at the time of launch, the reminder literally highlighted "dismiss" as an option—"over the user exiting the platform." Ex. 987 ¶ 206. This design employs the dark patterns known as "interface interference" and "visual prominence," *id.*, encouraging users to dismiss their reminders and continue using YouTube. *See* Ex. 1033 at 0060.

Time Watched Profile. Time Watched Profile was aimed at heavy habitual use. *See* Ex. 1034 at 9500. It is also an opt-in tool, meaning it is subject to the same friction described above. Even if accessed, it is a passive tool that simply allows users to check their average watch time. *See id.* The tool provides no interruption, suggestion, prompt, or guidance regarding continued use of the platform.

Notification Clusters. Notification Clusters allowed users to modify when they received notifications. *See* Ex. 1033 at 060. At launch, it was an opt-in setting for users to bundle push notifications into a single daily notification and set a specific time to receive a notification digest. *See* Ex. 799 at 5966. Again, it was a hard-to-find opt-in tool, ensuring few would find out about it and turn it on. *See* Ex. 1033 at 0060. In June 2018, YouTube noted that [redacted] users had enabled the notifications schedule digest. *See* Ex. 692 at 9739. In 2018, YouTube had over a billion active users, meaning that the opt-in rate for notifications clusters was just [redacted]. This is not surprising given YouTube's "success criteria"—ensuring the feature had a "neutral" impact on watch time. Ex. 777 at 6605.

Bedtime Reminders. This tool was launched to address late night use. *See* Ex. 730 at 5; Ex. 1034 at 9501. Users who opt-in can set specific times to be reminded to stop watching videos and go to bed. *See* Ex. 1034 at 9501. As with the other tools described above, Bedtime Reminders was hard to find and, even if enabled, easy to dismiss. The reminder appears with two options, "dismiss" or "snooze," and content continues to be displayed. Ex. 1063 at 17, 22. There is no option to immediately leave the YouTube platform. Further, Bedtime Reminders was only available on certain devices and, up until 2022, not available for Shorts (despite Shorts being YouTube's flagship teen-oriented product). *See* Ex. 730 at 5.

Take a Break Reminders, Time Watched Profile, Notification Clusters, and Bedtime Reminders are also all self-policing tools, allowing users to set their own guardrails (or none at all). They are, therefore, unlikely to have a meaningful impact on young users, who YouTube knows have poor self-regulation and impulse control. Ex. 1037 at 41 (recognizing that teens and tweens are in a unique developmental stage and "Self-regulation is not as developed in their prefrontal cortex, so they are less able to moderate and self-reflect in the moment.").

In 2021, YouTube changed the settings for Take a Break and Bedtime Reminders, turning them *on* by default for "all users 13-17." Ex. 1064 at 0252. This was a modest step in the right direction. However, it did not change the fundamental limitation of these tools—they place the onus on young users to self-police and self-regulate, by actually honoring the reminders they're given. As any parent of a tween or teen knows firsthand, "although more developed than children, adolescents are still developing self-regulation, meaning that they may still struggle to control impulses and inhibit urges." Ex. 983 ¶ 47. YouTube tapped into this aspect of adolescent psychology by encouraging teen users to "dismiss" or

"snooze" the reminders provided.

YouTube has been happy to publicly tout these digital wellbeing tools, but it is far less focused on their efficacy. Indeed, YouTube does not track whether users act on breaks or bedtime reminders—one witness testified, "That is unknowable" and "We don't track that." Ex. 1031 at 57:7-13, 59:9-20. (As discussed above, this is not "unknowable" for Meta, which does track adoption rates for its comparable screen time management tools. Those rates confirm that opt-in tools are very rarely used. *See supra* § III.A.2.a.c.) One employee who inquired about the success of YouTube's digital wellbeing tools was sent links tallying the number of users who *saw* the reminders—but no meaningful metric of their actual effectiveness in limiting use. Ex. 777 at 6605; *see also* Ex. 779 at 31; Ex. 715 at 457:25-458:23; Ex. 1022 at 156:3-20; 254:3-255:23; Ex. 1031 at 58:12-19, 59:9-17 (did not track if people actually turned off YouTube app after receiving a take a break or bedtime reminder). But YouTube continues to parade these tools before users, their caregivers, and teachers, without warning them that there is no reason to believe they meaningfully help users reduce screen time.

Tellingly, YouTube *does* track the impact of its screen time management tools on its bottom-line business metrics—[redacted] per Ex. 778 at 11) and [redacted]. As one document concedes: "We also check [redacted] when rolling out new features to make sure we didn't break anything. They're expected to be neutral though." Ex. 777 at 6605. The idea that screen time management tools would have a [redacted] is counterintuitive, to say the least. Ex. 982 ¶ 177-178. And it suggests why YouTube declined to measure the efficacy of these tools—it knew they were useless.

(i) Beauty Filters

YouTube provides users with augmented reality (AR) effects and filters that users can apply to modify their videos. Ex. 1013 at 73:4-25. This includes filters that can change backgrounds and apply fun masks. It has also included filters that can change a user's appearance—by adjusting eye color, brightening teeth, smoothing skin texture, and changing skin tones. *Id.* at 73:4-13. YouTube launched these filters because it knows they are particularly attractive to young users. Ex. 1055 at 9400-01 (email discussing TikTok's "Bold glamour" filter attracting "millions"); Ex. 1065 at 6687-88. It also knows these filters "drive video creation and consumption." Ex. 1072 at 6050 (funding Effects team to accelerate this); *see*

also Ex. 701 at 1277 (2018: "If Snapchat and Instagram are any indication, creation follows new effects."). In 2021, YouTube introduced these appearance-altering effects to Shorts, Ex. 693 at 2908, despite knowing it is primarily used by adolescents, Ex. 844 at 4694; Ex. 858 at 3; Ex. 733 at 8317 ("Shorts is our big thing for teen appeal"). As recently as 2024, YouTube was exploring how it could expand visual effects, including those primarily focused on appearance enhancements. *See* Ex. 1057 at 1837, 1849.

YouTube keeps these effects and filters on its platform despite knowing they can be dangerous to youth mental health. Jessica DiVento Dzuban, formerly YouTube's Global Head of Mental Health (and now TikTok's Global Head of Mental Health Policy) testified that she spoke with numerous youth advocacy groups, including the American Academy of Pediatrics, regarding parents' concerns that YouTube was affecting their child's mental health. Ex. 1024 at 81:9-82:6. Dr. Dzuban specifically recalled the discussion centering on body image and exposure to instructions on how to engage in risky behavior. *Id*. Despite those conversations, YouTube provides beauty filters for use by minors, Ex. 1013 at 73:15-74:13, and allows filters created by third parties, *id.* at 75:14-17.

YouTube knows that appearance altering filters encourage unrealistic and harmful beauty standards and perpetuate cultural and racial biases. *See, e.g.*, Ex. 1053 at 5593 (encouraging unrealistic standards); Ex. 1065 at 6687 (perpetuating bias); Ex. 1055 at 9400 (advancing unhealthy ideals). Indeed, its own research established the profound psychological dangers of these tools. *See* Ex. 1073 at 31-32 (July 2023 internal presentation identifying specific teen wellbeing concerns pre-dating facial deformation effects and noting, with increasingly sophisticated technology, "it is reasonable to predict negative outcomes in line with or potentially worse than what we've previously seen."); *see, e.g.*, Ex. 1074 at 7 ("expert researchers found beauty filters and negative social comparison may be correlated with negative wellbeing, especially in youth" and "the pervasive nature of filtered images regularly trigger body dysmorphia.").

One YouTube user research study from July 2023 documented that "participants talked at length about how facial deformation effects" made them feel worse about their natural appearance." Ex. 1073 at 33; *see id.* at 29 (deformation effects are those that can only be replicated in real life through surgical intervention, if at all). Users reported not only individual psychological harm, but "social stigma and shame that accompanies the use of facial deformation effects." *Id*. at 34. YouTube was clear-eyed

about how these effects could result in social anxiety and isolation, describing the following "stages of impact": "I cannot meet people in real life and disappoint them," "I cannot show my real self anymore," "I need to also look this way," and "everyone's perfect but me." Ex. 1075 at 6913. All of this is aligned with the Districts' experts' testimony about beauty filters. *See e.g.,* Ex. 982 ¶ 195 (because of filters, "users are now exposed to 'manipulated' photos that depict unrealistic (and in some cases impossible) body image standards"); Ex. 983 ¶ 5.f ("Filters and other image manipulation options take advantage of pubertal development, adolescent egocentrism, and social development"); Ex. 1053 at 5593-95.

YouTube's decision to allow adolescents to utilize appearance-altering beauty filters, despite this research, is a cynical prioritization of profits over child safety.

(6) YouTube's misrepresentations and manipulations

On August 10, 2021, YouTube publicly announced the launch of "new safety and digital wellbeing options for younger people on YouTube," including certain of the screen time management tools discussed above. Ex. 872 at 1. In a blog post, YouTube touted its "ongoing efforts to create age-appropriate experiences for young people on YouTube," observing "Young people are our future." *Id* at 4. Unfortunately, while YouTube has been public and vocal about its safety features (inadequate and lackluster as they are), behind closed doors it has undermined those features by ramping up efforts to increase youth consumption. For example, as the same time YouTube was announcing the launch of Bedtime Reminders, it was also finalizing the design of an infinite scroll experience on mobile Watch Next. Ex. 792 at 5068. At the same time YouTube offered tools to reduce distraction by "cluster[ing] notifications," it was setting a goal to *increase* the total number of notifications sent. Ex. 1040 at 9; Ex. 1048 at Growth Tab, Row 4. YouTube's well-being tools could not and did not solve the fundamental problem that the platform is designed to reduce rational decision making and optimize for more Watch Time. Ex. 1041 at 27.

Moreover, YouTube's announcements about keeping safe "users ages 13-17 on YouTube," Ex. 872 at 2; Ex. 761 at 2 (PTA editorial); Ex. 702 at 51:8-13; blow past the fact that YouTube has no idea who those users are—and has deliberately avoided finding out. Ex. 1012 at 105:19-24 (only a small fraction of those under 18 accurately declare their ages). YouTube could have done better, as it already was outside the United States. *Id.* at 143:22-144:23.

(7) YouTube's failure to warn

YouTube failed to warn parents, the Districts, and their students about the dangers caused by the YouTube platform to young users. It knew that school administrators and parents wanted more effective time management controls, to minimize distraction for students and enforce responsible use during the school day. Ex. 1062 at 0828-29; Ex. 766 at 462:14-469:22; 470:2-20, 476:21-480:3, 471:3-472:22; Ex. 764 at 2982-83. It knew that students were easily able to bypass existing tools like MRM, yet it failed to alert schools, including the Districts, to this vulnerability. *See* Ex. 740 at 250:6-11; 252:16-256:3, 282:2-283:1.

Meanwhile, YouTube knew that its persistent focus on increasing Watch Time was having the predictable and intended effect of keeping teens on its platform late into the night—driving a loss of sleep and, from there, distraction, loss of productivity, anxiety, and depression. Ex. 1003 at 227:6-9, 230:13-19; Ex. 1058 at 2685 (data reporting that approximately 7 percent of teens watch YouTube past midnight on school nights, and up to 30 percent note loss of sleep); Ex. 1059 at 33 (finding most people need 7 to 9 hours of sleep to be productive but blue light blocks the brain's signals to sleep); Ex. 738 at 15-17, 35, 40, 42; Ex. 774 at 14-25 ("disabling or limiting Autoplay during the night could result in sleep savings"); Ex. 775 at 6; Ex. 1058 at 2685 ("~7% of teens on YT watch past midnight on school nights. Teen 'night owls' were 88% more likely to have emotional & behavioral problems....30% of users 18-24 say YouTube has cut into sleep"). And while it knew its heavily advertised "Take a Break" and "Bedtime" reminders were minimally effective, it said nothing. Ex. 1061 at 43; Ex. 732 at 9261.

It is reasonably foreseeable that YouTube's failure to warn would harm the Districts, by requiring them to divert limited resources to manage YouTube-created problems.

3. Defendants' conduct substantially interferes with the public right to health, safety, and education in Arizona, Georgia, Kentucky, and Maryland

In the preceding pages, the Districts have illustrated the scale of Defendants' wrongdoing. Defendants' conduct breaches their duty of care to the Districts and is legally negligent. It also constitutes a significant interference with the public's right to health, safety, and education. The Court has permitted Breathitt, Tucson, DeKalb, and Harford to seek remedies for that interference through a public nuisance cause of action. *See In re Soc. Media Adolescent Addiction/Pers. Inj. Prods. Liab. Litig.*, 2024 WL

5694360, at *17 (N.D. Cal. Nov. 15, 2024) ("Public Nuisance MTD Order").

Defendants argue in a cursory fashion that these Districts have not demonstrated a *significant* interference with a public right. Defs.' Breathitt Mot. at 12-13 ("Breathitt must come forward with evidence that each Defendant's actionable conduct was a direct cause of a *significant* interference with a public right…"); Defs.' Dekalb Mot. at 12-14 (similar); Defs.' Tucson Mot. at 15-17 (similar); Defs.' Harford Mot. at 15-16 (similar). As an initial matter. Defendants waived this argument by failing to develop it. *Beasley v. Astrue*, 2011 WL 1327130, at *2 (W.D. Wash. Mar. 24, 2011) (cursorily raising an issue without any supporting analysis or explanation waives the issue raised). Defendants' argument also fails on the merits. Defendants cite no cases granting summary judgment on the issue of whether an interference with public rights is "significant." That is unsurprising, because the significance of a defendant's contribution to a public nuisance is a fact-intensive "question properly left to the jury." *In re Nat'l Prescrip. Opiate Litig.*, 589 F. Supp. 3d 790, 809 (N.D. Ohio 2022).

In the preceding pages of this omnibus brief, and in each District's accompanying submission, each District has adduced evidence that the design of Defendants' platforms has led millions of young people to experience compulsive use and attendant mental health issues, which significantly disrupts and harms school operations and the school environment, and frustrates the ability of schools to achieve their mandate of educating students. Each District has also put forth evidence that Defendants' conduct foreseeably caused it to expend and divert resources to address these problems. *See, e.g.*, Ex. 1000 ¶¶ 5, 123. This is precisely the type of evidence this Court found may support a significant interference with the public right to health and safety. MTD Order, 754 F. Supp. 3d at 956-59; Public Nuisance Order, 2024 WL 5694360, at *14 & n.20. Moreover, while the Court declined to address the public right to education at the pleading stage, the evidence set forth in this submission and the Districts' accompanying submissions shows that Defendants' conduct violated the public right to education in Arizona, Georgia, Kentucky, and Maryland. Public Nuisance MTD Order, 2024 WL 5694360, at *14. All four states recognize a public right to education. Az. Const. art. XI, § 1 *et seq*.; Ga. Const. art. VIII, § 1 *et seq*.; Ky. Const. § 183; Md. Const. art. VIII, § 1 *et seq*.

This evidence is sufficient for a jury to determine there has been a "significant interference" in a public right and accordingly, creates a triable issue of fact under Arizona, Georgia, Kentucky, and

Maryland nuisance law. *Nolan v. Starlight Pines Homeowners Ass'n*, 216 Ariz. 482, 489 (Ct. App. 2007) (significant interference means something "real and appreciable" that is "more than slight inconvenience or petty annoyance") (citing Restatement (Second) of Torts § 821F); *see also Blondell v. Courtney Station 300 LLC*, 865 S.E.2d 589, 601 (Ga. Ct. App. 2021).

Given that the Districts have proffered evidence of Defendants' significant interference with public rights to health, safety, and education, Defendants are not entitled to summary adjudication on this issue.

B. Defendants' causation arguments are premised on disputed facts and legal errors.

1. Defendants improperly raise the evidentiary bar in requiring Plaintiffs to quantify specific student usage

In each bellwether state, questions of actual and proximate causation are for the jury to decide, unless the evidence supports only one inference, such that no reasonable juror could find for the plaintiff. *Torres v. Jai Dining Servs. (Phoenix) Inc.*, 497 P.3d 481, 486 (Ariz. 2021); *Johnson v. Wood*, 2025 WL 2715888, at *4 (Ga. Ct. App. Sept. 24, 2025); *Pathways, Inc. v. Hammons*, 113 S.W.3d 85, 89 (Ky. 2003); *Pittway Corp. v. Collins*, 973 A.2d 771, 253 (Md. Ct. App. 2009); *Gilbert v. Stewart*, 255 A.3d 1101, 1113-14 (N.J. 2021); *Wickersham v. Ford Motor Co.*, 853 S.E.2d 329, 390–91 (S.C. 2020). "Only in plain and undisputed cases should a court even consider resolving such fact-intensive questions as proximate causation on summary judgment." *Meraz v. Ford Motor Co.*, 2014 WL 12558123, at *12 (C.D. Cal. June 13, 2014) (cleaned up). Defendants have not come close to meeting this high standard.

Defendants argue that causation can only be demonstrated if the Districts can precisely quantify the exact harm caused by each Defendant by furnishing data regarding individual student use of each individual feature of each individual platform. Defs.' Tucson Mot. at 5–7, 14–15; Defs.' Dekalb Mot. at 5, 12–13; Defs.' Breathitt Mot. at 12–13; Defs.' Harford Mot. at 14–15; Defs.' Irvington Mot. at 14–15; Defs.' Charleston Mot. at 14–15. This argument fails for several reasons.

First, Defendants' extreme position—that schools should be able to produce quantifiable evidence measuring their students' use by platform, duration, and even feature—does not reflect the law, and is not consistent with the facts of this case. It is for the jury to decide whether the Districts' evidence is sufficient to establish causation, not for Defendants to dictate the form of that that evidence. In *City & Cnty. of San Francisco v. Purdue Pharma L.P.*, the court rejected the argument that proof of causation required tracing

each defendant's individual prescription all the way through to the end-user's specific harm, because "the substantial factor test does not require 'the kind of incontrovertible linkage proposed by'" Defendants. 620 F. Supp. 3d 936, 1003 (N.D. Cal. 2022). The court found "it is difficult to take seriously the argument that the failure to identify specific" such data "is fatal to Plaintiff's case." *Id.* at 1004, 1004 n.31. So too here. First, there are legal constraints on schools' ability to obtain student usage data. Moreover, as a practical matter, schools are in the business of teaching and caring for students, not meticulously cataloging each instance of social media usage. As Dr. Hoover has explained, school districts are resource-constrained institutions and "existing [school district data] systems often do not track the types of internal service referrals, behavioral disruptions, or emotional dysregulation commonly tied to social media use." Ex. 1103 ¶ 14; Ex. 1102 ¶14; Ex. 1101 ¶14; Ex. 1104 ¶ 14; Ex. 1105 ¶ 14; Ex. 1100 ¶ 16. In a telling contradiction, Defendants oppose the remedy of better data infrastructure for the Districts to track such events. *See* Defs.' Tuscon Mot. at 11; Defs.' Deklab Mot. at 8–9; Defs.' Breathitt Mot. at 9; Defs.' Harford Mot. at 10; Defs.' Irvington Mot. at 9; Defs.' Charleston Mot. at 9–10.

Second, Defendants' argument ignores the substantial evidence the Districts have produced through fact and expert testimony, showing that each District was foreseeably harmed by widespread, problematic use of Defendants' platforms. *See supra* § III.A.2; *see* Ex. 1000 ¶ 30; Ex. 656 ¶¶ 54-75; Ex. 982 at ¶¶ 269-530; Ex. 1004 ¶¶ 311-50; Ex. 1005 ¶¶ 5.a-5.j; Ex. 981 ¶¶ 282-367; Ex. 1095 ¶¶ 48-52; Ex. 1096 ¶¶ 49-51; Ex. 1093 ¶¶ 48-52; Ex. 1091 ¶¶ 52-54; Ex. 1092 ¶¶ 46-50; Ex. 1094 ¶¶ 51-55; Ex. 656 ¶¶ 54-75. This evidence incorporates a broad array of facts that support the schools' claims and is more than sufficient to create disputed issues of material fact. *In re Nat'l Prescription Opiate Litig.*, 2019 WL 4178617, at *3 (N.D. Ohio Sept. 3, 2019) (denying summary judgment and rejecting argument that plaintiffs could not prove causation against multiple defendants who all contributed to injuries).

Third, Defendants are mistaken in arguing that references to "social media" by administrators, teachers, and counselors are not probative. Defendants are not obscure, niche platforms—they are the most popular social media platforms used by American youth. Ex. 981 ¶¶ 284-87, 290, 329-30; Ex. 1038 at 1-2. A reasonable jury could infer that when educators and students refer to "social media," they mean Instagram, TikTok, YouTube, and Snapchat—even if each reference does not name the platforms explicitly. *See* Ex. 1000 ¶ 30 (Dr. Hoover explaining that "[a]ccording to a recent Pew Research Center

report, over a third of teens say they are using the top five online platforms (YouTube, TikTok, Instagram, Snapchat and Facebook) almost constantly"); Ex. 1097A at 203:7-24 ("[S]ince you've asked about the four defendants' platforms, those are … by the work that I do with students and schools, the ones that are most often brought up and used … When I speak about social media, [in my reports] it's talking about the platforms that young people are using, and those happen to be the defendants' platforms."). Even if these statements were viewed as "circumstantial evidence," that is enough: "California courts have repeatedly held that causation may be reasonably inferred from circumstantial evidence in the context of mass torts." *City & Cnty. of San Francisco v. Purdue Pharma L.P.*, 620 F. Supp. 3d 936, 1002 (N.D. Cal. 2022). The same principle applies under the laws of every bellwether state.[8] And at summary judgment, all inferences must be drawn in favor of the Districts. *Hunt v. City of L.A.*, 638 F.3d 703, 709 (9th Cir. 2011).

Fourth, Defendants are wrong that the Districts must produce evidence apportioning a precise percentage of liability to each Defendant. Determining the relative significance of each Defendant's conduct is a question for the jury. *See In re Nat'l Prescription Opiate Litig.*, 589 F. Supp. 3d 790, 808 (N.D. Ohio 2022) ("Determining the significance of each Defendant's . . . conduct [in causing the nuisance] was a question properly left to the jury); *see also Social Media Cases*, 2025 WL 2807828, at *7 (Cal. Super. Sep. 22, 2025) ("Determining whether the substance or instrumentality upon which liability is premised caused a particular plaintiff's harm is . . . the role of the jury."). A nearly identical argument was rejected in the Opioids MDL. There, a network of manufacturers, distributors, and pharmacies each contributed to a nationwide public nuisance. The defendants argued that the local government plaintiffs could not prove the causal contribution of any individual defendant to any particular county without individualized evidence. *In re Nat'l Prescription Opiate Litig.*, 589 F. Supp. 3d at 808. The court squarely

[8] *Schall v. Suzuki Motor of Am., Inc.*, 450 F. Supp. 3d 771, 784 (W.D. Ky. 2020) ("Causation may be proved by circumstantial evidence"); *Mason v. Arizona Pub. Serv. Co.*, 622 P.2d 493, 500 (Ariz. Ct. App. 1980) ("Proximate cause may be determined from circumstantial evidence."); *J.T. Baggerly v. CSX Transp., Inc.*, 635 S.E.2d 97, 101 (S.C. 2006) ("Normally, proximate cause is a question of fact for the jury, and it may be proved by direct or circumstantial evidence."); *Furlong v. Dyal*, 539 S.E.2d 836, 838 (Ga. Ct. App. 2000) (plaintiff permitted to rely on "circumstantial and opinion evidence to create a jury question" "on the issue of causation"); *Rowhouses, Inc. v. Smith*, 133 A.3d 1054, 1067 (Md. 2016) (circumstantial evidence and expert testimony can suffice to prove causation); *Underwood v. Camden Cnty. Off. of Sheriff*, 2024 WL 1366672, at *3 (D.N.J. Mar. 28, 2024) ("circumstantial evidence can be used to establish causation").

rejected that argument, holding it "ignore[d] the aggregate nature of the evidence presented at trial and the natural inferences allowed therefrom." *Id.* The aggregate evidence "supported a reasonable inference that [each] Defendants' conduct was a substantial factor in creating the alleged nuisance." *Id.* The court concluded, "aggregate proof of causation [is] sufficient to overcome summary judgment." *In re Nat'l Prescription Opiate Litig.*, 2019 WL 4178617, at *3.

Similarly, the Districts have offered aggregate evidence showing that each Defendant's conduct was a substantial, concurring, or contributing factor in causing the Districts' harm. That is all the law requires. *See Bailey v. N. Am. Refractories Co.,* 95 S.W.3d 868, 871-73 (Ky. Ct. App. 2001) (applying substantial factor test under Kentucky law); *Stollenwerk v. Tri-W. Health Care All.*, 254 F. App'x 664 (9th Cir. 2007) (applying substantial factor test under Arizona law); *Gilbert v. Stewart,* 255 A.3d 1101, 1114 (N.J. 2021) (applying substantial factor test under New Jersey law); *Yonce v. SmithKline Beecham Clinical Lab'ys, Inc.*, 680 A.2d 569, 576-77 (Md. Ct. Sp. App. 1996) (applying substantial factor test under Maryland law); *J.T. Baggerly*, 635 S.E.2d at 101 (explaining that causation is satisfied where the evidence establishes that the defendant's negligence is "a concurring or a contributing proximate cause.") (citation omitted); *Hayes v. Crawford*, 730 S.E.2d 26, 29 (Ga. Ct. App. 2012) (similarly providing for concurrent proximate causation).

2. Defendants' "entanglement" arguments on causation are premised on legal errors

a) Defendants misapprehend the features at issue by writing this Court's failure to warn ruling out of existence.

To prevail on their failure to warn claims, the Districts are not required to prove that each individual platform feature independently caused or contributed to their harm. This Court has repeatedly held that Defendants may be held liable for failing to warn "of known risks of addiction attendant to any platform features ***or as to platform construction in general***." MTD Order, 754 F. Supp. 3d at 963–64 (emphasis added); *accord In re Soc. Media Adolescent Addiction/Pers. Inj. Prods. Liab. Litig.*, 753 F. Supp. 3d 849, 865 (N.D. Cal. 2024) ("AG MTD Order"). Defendants' suggestion that summary judgment is warranted because the District have not linked each specific feature to the harms alleged ignores—and would nullify—those rulings. In any event, the record contains ample evidence trying particular platform features to the harms suffered. *See supra* at § III.A.2.

b) The Districts need not rule out every possible alternative causal factor to avoid Section 230.

Similarly, Section 230 also does not bar the Districts' claims simply because their harms are factually related to "third-party wrongdoing or the publication of content." Mot. 16. Defendants' attempt to graft a "but-for" causation test onto Section 230 has been squarely rejected by the Ninth Circuit.

Defendants misread *Lemmon v. Snap, Inc.*, which held that "[t]he duty to design a reasonably safe product is fully independent of Snap's role in monitoring or publishing third-party content." 995 F.3d 1085, 1093 (9th Cir. 2021). In holding Section 230 did ***not*** bar claims premised on Snap's product-designer duties, the Ninth Circuit expressly acknowledged that "publishing content is a but-for cause of just about everything Snap is involved in," yet clarified that this fact "does not mean that" holding Snap liable for harm caused by its speed filter "seeks to hold Snap responsible in its capacity as a 'publisher or speaker.'" *Id.* Even though content may be involved in Snap's operations—including, of course, the speed filter, which was created to apply to users' person-to-person snaps—"that does not mean that" holding Snap liable for harm caused by the speed filter would be holding it responsible as a publisher. *Id.* Section 230 protection only attaches where a plaintiff's claim would require Snap to "alter[] the content that Snapchat's users generate" which is not being requested here. *Id.* at 1092; *see id.* ("Snap's alleged duty in this case thus 'has nothing to do with' its editing, monitoring, or removing of the content that its users generate through Snapchat.").

The rule from *Lemmon* is clear: under Section 230, a defendant may be held liable for its own design decisions—even when those decisions affect how it distributes third-party content—but it cannot be held liable for publishing or failing to remove that content. Yet in Defendants' telling, *Lemmon* imposes an unprecedented limitation on tort liability, allowing claims to proceed only if "fully independent" of any publishing activity. That is not what *Lemmon* said. The Ninth Circuit simply recognized that social media companies have a duty to design safe platforms—separate from their editorial choices about what content to publish. It did not hold that content must be "fully independent" of the causal chain. Indeed, if Defendants' selective reading were correct, *Lemmon* itself would have been dismissed—since the alleged harm there arose from Snap's Speed Filter, a feature inherently intertwined with user content. *See* 995

F.3d at 1092 (recognizing that "Snap acted as the 'publisher or speaker' of user content by transmitting Landen's snap"). That is plainly not what *Lemmon* held, no matter how Defendants try to spin it.

At any rate, Defendants' attempt to reintroduce a "but for" test into *Lemmon* has been rejected by subsequent Ninth Circuit precedent. The court has clarified that Section 230 applies only when a plaintiff seeks to hold a defendant liable for third-party content, not when a claim challenges the company's own design choices, even if those choices relate to such content. In *Bride v. Yolo Technologies, Inc.*, the court explained that the claim in *Lemmon* "did not depend on third-party content"—and Section 230 therefore did not apply—because it "did 'not depend on what messages, if any, a Snapchat user employing the Speed Filter actually sends.'" 112 F.4th 1168, 1180 (9th Cir. 2024) (quoting *Lemmon*, 995 F.3d at 1094). In *Doe 1 v. Twitter*, 148 F.4th 635 (9th Cir. 2025), the court reaffirmed that *Lemmon* involved treating Snap as a "'product designer' rather than a 'publisher or speaker,'" because the claim "turned on Snap's design architecture rather than the publication of any content." *Id.* at 645. The *Doe 1* plaintiffs alleged defects in Twitter's mechanism for reporting CSAM—defects inseparable from the existence of third-party content. Yet Section 230 still did not apply, because the claim targeted Twitter's defective "reporting mechanism," which could be fixed "without monitoring, removing, or in any way engaging with third-party content." *Id.*

Calise v. Meta Platforms, Inc. underscores the same principle. In that case, the Ninth Circuit explained, "it is not enough that a claim, **including its underlying facts**, stems from third-party content for § 230 immunity to apply." 103 F.4th 732, 742 (9th Cir. 2024) (emphasis added). Indeed, the court recognized that nearly any claim alleging a failure to warn about an online platform could, in some sense, "concern content." *Id.* at 742. But that alone is not enough to trigger Section 230. As *Calise* observed, "it is not true that 'providing a warning' would not require 'considering the content posted.' How could Internet Brands warn about certain harmful content without considering what it was? Such a rudimentary fact-bound inquiry quickly falls apart and runs up against our precedent." *Id.* Section 230 was never intended to bar claims simply because content exists in the background. Sometimes content is involved, and that's fine.

In the bellwether cases, third-party content is related to various defective design features of Defendants' platforms. But, importantly, the Districts do not allege harm from the content itself—and, as

school districts rather than individual users, they do not allege having viewed content ***at all***. Under these circumstances, Section 230 is plainly inapplicable.

The additional cases Defendants cite do not help them. In *M.P. ex rel. Pinckney v. Meta Platforms Inc.*, 127 F.4th 516 (4th Cir. 2025), the Fourth Circuit applied Section 230 because the plaintiff sought to "hold Facebook liable for disseminating 'improper content' on its website." *Id.* at 525. Likewise, in *Doe (K.B.) v. Backpage.com, LLC*, 724 F. Supp. 3d 882 (N.D. Cal. 2024), Section 230 applied because the plaintiff's "entire claim [was] based on the content of the subsequent interaction" with a third party." *Id.* at 885. *Nuveen Mun. High Income Opportunity Fund v. City of Alameda*, 730 F.3d 1111, 1121 (9th Cir. 2013) and *In re Williams Sec. Litig.-WCG Subclass*, 558 F.3d 1130, 1138 (10th Cir. 2009) addressed loss causation in securities fraud, not factual or proximate causation in tort. Here, by contrast, the record contains ample evidence creating triable issues of causation.

In reply, Defendants may also point to *Doe v. Grindr Inc.*, 128 F.4th 1148, 1153 (9th Cir. 2025), but that case does not support their expansive view of Section 230. There, the plaintiff challenged Grindr's decision to allow him "to communicate with abusive adults." *Id.* In *Grindr*, the communication itself was the harm—the wrongful act and injury were one and the same. The claim thus turned on third-party content and publishing activity; the alleged harm and content were inseparable, not "independent" of each other. *See id.* By contrast, the Districts' claims do not arise from any third-party communication or content, but from Defendants' own defective platform design and failure to warn of known risks. *See id.*

c) Juries are routinely called upon to evaluate liability in the context of concurrent causation

Defendants are also wrong as a matter of basic tort law: They are not entitled to summary judgment simply because the Districts cannot perfectly isolate the harm caused by each platform. Under settled common law, "a tortfeasor is liable for any injury of which his negligence is a proximate cause." *Am. Motorcycle Ass'n v. Super. Ct.*, 578 P.2d 899, 904 (Cal. 1978). Even if "it is simply impossible to determine whether or not a particular concurrent tortfeasor's negligence, acting alone, would have caused the same injury," each may still be held liable for the harm it proximately caused. *Id.* at 589.

Left unsaid in Defendants' motions is that every bellwether jurisdiction applies the same rule: proximate cause exists where a defendant's "conduct was a substantial factor in bringing about the injuries,

even where there are other intervening causes which were foreseeable or were normal incidents of the risk created." *Gilbert v. Stewart*, 255 A.3d 1101, 1114 (N.J. 2021). This standard "does not require an unsevered connecting link between the negligent conduct and the ultimate harm." *Id.* (emphasis omitted). Rather, a plaintiff need only show it is "reasonably probable" that each defendant's conduct contributed to its injuries. *E.g.*, *Walton v. Premier Soccer Club, Inc.*, 334 A.3d 784, 796–97 (Md. 2025).[9] The Districts survive summary judgment on this ground because they have established it is "more likely than not" each Defendant's platform contributed to their alleged harm. *Walton*, 334 A.3d at 793. It would defy logic and settled tort law to hold that each Defendant may be a "substantial" factor in causing harm, yet escape liability simply because their misconduct is intertwined with that of other tortfeasors.

These basic legal principles also resolve Defendants' argument that the Districts must completely isolate the harm caused by their defective features from any potential negative effects from third-party misconduct, or even other platform features Defendants claim are not at issue (despite this Court's failure-to-warn ruling). The existence of "preexisting social issues affecting the community" does not undermine the Districts' claims, and the Districts are not required "to precisely apportion [their] injuries" to the alleged harm rather than to unrelated societal factors. *Ramah Navajo Sch. Bd., Inc. v. Sebelius*, 2013 WL 12303945, at *19 (D.N.M. May 9, 2013). Instead, "it is the duty of the trier of fact to make the best estimate in its power to base its damages calculation on the actionable injury itself." *Id.*; *see* Restatement (Second) of Torts § 433A(1) cmt. b (where there are multiple concurrent causes, damages should be awarded if "it is possible to make a rough estimate which will fairly apportion such subsidiary elements of damages."); *see also, e.g.*, *CRS Sirrine, Inc. v. Dravo Corp.*, 464 S.E.2d 897, 900 (Ga. Ct. App. 1995)

[9] The other bellwether states' laws are in accord. *See, e.g.*, *Mayer v. Hous. Auth. of Jersey City*, 202 A.2d 439, 447 (N.J. Super. Ct. App. Div. 1964), *aff'd* 210 A.2d 617 (N.J. 1965) ("If the evidence, viewed in the light most favorable to plaintiff, would cause fair-minded men to differ as to whether there was a reasonably probable relation of cause and effect between the alleged negligence and the injuries, the issue must be submitted to the jury."); *Robertson v. Sixpence Inns of Am., Inc.*, 789 P.2d 1040, 1046–47 (Ariz. 1990) ("Plaintiff need only present probable facts from which the causal relationship reasonably may be inferred."); *Barrett Props., LLC v. Roberts Capitol, Inc.*, 729 S.E.2d 621, 624 (Ga. Ct. App. 2012) ("To establish proximate cause, a plaintiff….must introduce evidence which affords a reasonable basis for the conclusion that it is more likely than not that the conduct of the defendant was a cause in fact of the result."); *accord Thacker v. Ethicon, Inc.*, 47 F.4th 451, 460 (6th Cir. 2022); *Allen v. Greenville Hotel Partners, Inc.*, 2006 WL 1817804, at *6 (D.S.C. June 30, 2006) (quoting *Ballou v. Sigma Nu Gen. Fraternity*, 352 S.E.2d 488, 493 (S.C. Ct. App. 1986)) (In South Carolina, causation can only be resolved on a motion for summary judgment in "rare or exceptional cases.").

("if a plaintiff can show with reasonable certainty the total amount of damages and the degree to which those damages are attributable to defendant, that is sufficient to support an award."). Aside from failure-to-warn, evidence concerning features shielded by Section 230 remains relevant for context.[10]

C. Genuine material disputes of fact concerning the Districts' damages preclude summary judgment.

1. Each Plaintiff has presented evidence of compensable "hard costs."

The Court ruled that school districts may seek damages for various costs incurred as a result of Defendants' negligence and nuisance, including but not limited to: "[1] diverting and increasing financial resources to address the disruptive forces of defendants' social media products in school; [2] hiring mental health personnel and developing mental health resources; [3] implementing new information technology and physical resources to limit access to and mitigate risks caused by defendants' platforms; and [4] repairing property damage." MTD Order, 754 F. Supp. 3d at 967. Consistent with the Court's ruling, the Districts have supplied invoices detailing their costs, which were described in the expert reports of Jeffrey Meyers. *See* Exs. 574, 575, 576, 577, 578, 579. District employees have likewise attested that these costs directly resulted from Defendants' misconduct. *See infra* § III.C.1.a.

Defendants do not claim that the Districts' hard costs are not recoverable as a matter of law and, with limited exceptions, do not challenge whether the Districts have evidence of those costs. Instead, they claim that the schools have no "competent" evidence of past damages. Defs.' Tucson Mot. at 25; Defs.' Dekalb Mot. at 22; Defs.' Breathitt Mot. at 20; Defs.' Irvington Mot. at 21; Defs.'Charleston Mot. at 23. This is shorthand for asking the Court to cut out the jury and *weigh* the evidence, which is impermissible at summary judgment. *See Berman v. Freedom Fin. Network, LLC*, 400 F. Supp. 3d 964, 969 (N.D. Cal. 2019) (Gonzalez-Rogers, J.) ("the weighing of the evidence, and the drawing of legitimate inferences from facts are jury functions, not those of a judge.").

Because the schools have presented evidence that their costs were incurred due to Defendants' misconduct, the question is for the jury to weigh and make an ultimate finding about causation. In every bellwether state, questions of general and proximate cause are for the jury unless there is *no* possibility of

[10] The Districts more fully respond to this misguided argument in the 230 Daubert Opp.

an inference supporting the non-moving party. *Torres v. Jai Dining Servs. (Phoenix) Inc.*, 497 P.3d 481, 486 (Ariz. 2021); *Johnson v. Wood*, 2025 WL 2715888, at *4 (Ga. Ct. App. Sept. 24, 2025) ("Questions of negligence, diligence, contributory negligence, and proximate cause are peculiarly matters for the jury, and a court should not take the place of the jury in solving them, except in plain and indisputable cases." (citation omitted)); *Pathways, Inc. v. Hammons*, 113 S.W.3d 85, 89 (Ky. 2003); *Pittway Corp. v. Collins*, 973 A.2d 771, 253 (Md. Ct. App. 2009); *Gilbert v. Stewart*, 255 A.3d 1101, 1113–14 (N.J. 2021); *Wickersham v. Ford Motor Co.*, 853 S.E.2d 329, 390–91 (S.C. 2020). This is not one of those rare cases. The record contains ample evidence of causation, which remains disputed and must therefore be resolved by the jury. Defendants can argue to the jury that these costs were attributable to some other cause, and the jury will evaluate and weigh the evidence presented.

a) The Districts presented evidence of incurred costs relating to mental health resources, and technology materials, systems, and programming.

The Districts have adduced evidence of costs incurred to address Defendants' misconduct, such as costs of adding mental health resources and implementing new technology, including physical materials, systems, and programming. The Court has already confirmed that such categories of damages are recoverable. MTD Order, 754 F. Supp. 3d at 963.

Technology Lockers. Several Districts seek compensation for purchasing physical technology lockers (i.e. Yondr pouches[11], Amazon Capital lockers, or other similar devices) in response to Defendants' conduct.[12] While Defendants object to these costs by arguing the lockers are intended to stop students "from accessing their phones *generally*, not just Defendants' platforms," the Districts have

[11] Yondr pouches are locking magnetic pouches purchased by schools to secure students' phones during the school day, limit cell phone and social media use, reduce classroom distractions, and promote a more focused learning environment.

[12] **Technology Lockers. Tucson**: Yondr pouches ($16,850) and Amazon Capital lockers ($16,101). *See* Ex. 579 at App'x C Ex. 1 at 1-2. **DeKalb**: Yondr pouches ($391,076) and McGee phone lockers ($21,280). *See* Ex. 576 at App'x C Ex. 1 at 1. **Breathitt**: Quill cellphone caddies ($209) and Amazon Capital lockers ($1,758). *See* Ex. 574 at App'x C Ex. 1 at 1. **Charleston**: Yondr pouches ($273,953). *See* Ex. 575 at App'x C Ex. 1 at 1. **Harford** and **Irvington** do not seek compensation for technology lockers.

presented evidence that the lockers would not have been purchased but for social media use in schools.[13] Defendants do not engage with these proffered facts, which create a clear question for the jury to determine whether social media use from Defendants' platforms was a substantial factor in requiring the purchase of technology lockers. That there may be other benefits of locking phones away during school is not dispositive, and Defendants present no evidence that other cellphone use is the sole or superseding cause necessitating locker purchases.

Filtering and Monitoring Software. Certain Districts seek partial compensation for purchases of internet filter systems and school software that block social media websites in schools or assist school staff in monitoring students' use:

- **Breathitt**: 75% compensation for implementing monitoring software Hapara ($47,468); 50% compensation for implementing E-Hallpass software Eduspire ($725); and 25% compensation for implementing E-Hallpass software Securly ($2,222). *See* Ex. 574 at App'x C Ex. 1 at 1-3.
- **Tucson**: 40% compensation for implementing online reporting system Awareity ($33,582). *See* Ex. 579 at App'x C Ex. 1 at 1.
- **Irvington**: 35% compensation for implementing filtering systems GoGuardian ($73,185) and Palo Alto ($47,565). *See* Ex. 578 at App'x C Ex. 1 at 5-6.
- **DeKalb**: 30% compensation for implementing filtering system Lightspeed ($1,173,042). *See* Ex. 576 at App'x C Ex. 1 at 1.
- **Harford**: 20% compensation for implementing filtering systems Skyline ($68,447) and Cisco Umbrella ($64,346). *See* Ex. 577 at App'x C Ex. 1 at 1.
- **Charleston** does not seek compensation for internet filtering systems or monitoring software.

Defendants argue that these systems are not used *solely* to block or respond to social media, (*e.g.* Defs.' Tuscon Mot. at 34; Defs.' Breathitt Mot. at 25), or that schools may have been required to implement filtering software irrespective of Defendants' platforms (Defs.' Dekalb Mot. at 28; Defs.' Harford Mot. at 31-32; Defs.' Irvington Mot. at 27). However, the Districts have presented evidence that these costs are

[13] **Tucson**: Ex. 1125 at 41:19-44:3; Ex. 1126 ¶ 12 (Sabino installed cellphone lockers to physically remove phones and access to social media from students because alerts prompt checking social media instead of focusing on class); **DeKalb**: Ex. 1127 at 82:10-83:6 (implementation of pouches and lockers was a direct response to distractions from social media platforms); **Breathitt**: Ex. 1128 at 16:11-15 (district would not have purchased cell phone caddies "if it was not for social media"); **Charleston**: Ex. 1129 at 58:1-22.

in part attributable to Defendants' platforms, so these facts are in dispute.[14]

Mental Health Resources. Certain Districts seek compensation for new and increased mental health resources, as set out below. The District adduce evidence that Defendants' platforms were a substantial factor in necessitating expansion of these resources.[15]

- **Tucson**: 40% compensation for access to online therapy platform Talkspace ($231,021). *See* Ex. 579 at App'x C Ex. 1 at 2.
- **Irvington**: 20% compensation for mental health resources through New Jersey Coalition for Inclusive Education ($325,147), Care Plus NJ ($624,917), Brett Dinovi & Associates ($226,467), Live Breathe Calm ($19,770), Generations Family Guidance ($58,160), Center for Partnership Services ($146,136), and Momentum Therapy Services ($66,086). *See* Ex. 578 at App'x C Ex. 1 at 1-7.
- **Breathitt**, **Charleston**, **DeKalb**, and **Harford** and do not seek compensation for expanded mental health resources beyond the lost time for staff providing those resources, as discussed in the lost time section.

Defendants claim that the estimated percentage compensation for Tucson's resources lacks sufficient basis, *see* Defs.' Tucson Mot. at 34-35, and that one of Irvington's vendors charges a flat rate that it would have incurred absent Defendants' misconduct. *See* Defs.' Irvington Mot. at 27. *But see* Ex.579; Ex. 582; Ex. 668. These arguments go to the weight of the evidence and are for the jury to consider. Defendants admit that the District experts relied on sworn interrogatory responses that confirm these resources were required in part due to social media use in schools. The fact that these programs may provide benefit to some students beyond the context of social media use does not erase the fact that they were used in

[14] **Breathitt**: Ex. 661 at 22:11-23:9 (Hapara's focus is "keeping [students] off their social media and then also keeping them off videos, YouTube videos"), 23:16-24 (Hapara "was purchased above and beyond just the normal blocking and filters."); Ex. 1133 at 128:11–22 (Breathitt uses Hapara to disconnect students from social media); **DeKalb**: Ex. 1130 at 143:2–24 (Lightspeed is designed to block access to social media websites); Ex. 520 ¶¶ 8-15 ("30% of Lightspeed costs are related to the District's attempts to block and/or limit social media."); **Harford**: Ex. 598 at 28:6-25 (explaining 20% attribution of costs to social media events), 21:5-24:25 (describing filtering specific to social media, not just what law requires); Ex. 571 at 13:11 (describing that Cisco and Skyline are used to block social media).

[15] **Tucson**: Ex 543 at 72:12-17 (district has seen an "increase of anxiety and depression and overall mental health supports needed on [its] campuses," with conditions "heavily influenced by social media"), 73:4-16 (increased "anxiety, depression, and body dysmorphia" attributable to "social media platforms" drove need for increased mental health resources); **Irvington**: Ex. 524 at 99:22-100:8 (describing social media as a "main" reason why increased mental health resources were required); Ex. 600 at 146:1-147:16 (attributing social media as a substantial factor in mental health harms for IPS students, including by exacerbating preexisting problems).

response to Defendants' misconduct. Defendants offer no evidence that an alternate cause was the sole or superseding cause of these costs, and even if they did, that would be in dispute.

Programming. The same is true of evidence of new and increased social emotional learning (SEL) programming costs and other responsive costs, which the Districts have incurred as follows:

- **Breathitt**: 50% compensation for programming from Kids First ($498); 25% compensation for programming from Remix Education ($2,733); 50% compensation for programming from Millstone Labs ($6,719). *See* Ex. 574 at App'x C Ex. 1 at 1-2.
- **Tucson**: 40% compensation for programming with Character Strong ($242,856). *See* Ex. 579 at App'x C Ex. 1 at 1-2. 40% compensation for creating Comprehensive School Threat Assessment Guidelines ($2,912). *See* Ex. 579 at App'x C Ex. 1 at 2.
- **Charleston**: 7% compensation for programming from Flippen Group ($95,476); 20% compensation for implementing a social emotional learning platform from Panorama Education ($192,105); 40% compensation for restorative practices training and coaching from International Institute for Restorative Practices ($163,278); 35% compensation for social emotional learning curricula from Committee for Children for the Second Step ($226,278); 40% compensation for programming from Restorative Resolutions ($33,110); 40% compensation for educational training services provided by Restorative Coaching ($9,600); 80% compensation for educational services from Social Emotional Learning Alliance for South Carolina ($9,600). *See* Ex. 575 at App'x C Ex. 1 at 1-9.
- **DeKalb**, **Harford**, and **Irvington** are not seeking compensation for hard costs of expanded programming.

The Districts have presented evidence that they implemented various programming in part as a response to rising and disruptive social media use in schools.[16] It is for the jury to determine whether the evidence establishes that Defendants were a substantial factor in causing the Districts to incur those costs.

b) Defendants' cross-cutting entanglement arguments do not support summary judgment on causation

[16] **Tucson**: Ex. 543B at 354:10-15; *see also id.* at 357:21-358:6 (connecting the need for social emotional curriculum in Tucson to the negative impact of social media on students' social emotional well-being), 362:10-19 (noting that Character Strong curriculum is focused on social emotional skills and includes specific lessons about social media); Ex. 543A at 71:6-72:17 (increased anxiety, depression, and overall mental health supports needed by students heavily influenced by Defendants' platforms); Ex. 547 at 188:22-190:5 (identifying substantial work with SEL curriculum to help address students' needs resulting from student use of social media, specifically TikTok, Instagram, YouTube, and Snapchat); **Breathitt**: Ex. 1128 at 46:23–47:14 (Cyber Safe Week to discuss dangers of social media apps); **Charleston**: Ex. 548 ¶ 31 (Screenagers directly attributable to Ds' conduct); *id.* ¶ 26 (increased use of Panorama attributed to Ds' conduct); *id.* ¶ 24; *id.* ¶ 25 (connecting Flippen Group's Capturing Kids' Hearts to social media); Ex. 566B at 200:15-202:16 (apportioned for costs working with adults rather than students); *id.* at 64:15-95:22 (connecting Second Step to alleged harms); Ex. 1186 at 88:21-90:10 (connecting restorative practices curriculum to social media).

Defendants argue that the Districts failed to attribute their costs solely to Defendants' conduct. But this disregards a "long-recognized tenet of tort law" that there can be multiple "'but-for' causes" that support liability, and as a result a plaintiff is not required to show actionable conduct "was the 'sole' cause at summary judgment." *See Moore v. City of Atlanta, Ga.*, 2022 WL 19517299, at *25 (N.D. Ga. Dec. 19, 2022); *supra* § III.B.2.

Along the same lines, Defendants argue that the Districts' costs are too difficult to calculate or apportion given multiple potential causes. That is no basis to grant summary judgment on causation. Under the law of all bellwether states, apportionment of damages among defendants is a question for the jury that flows from its determination of fault for any (or all) of the Defendants. *Piner v. Sup. Ct. in and for Cnty. of Maricopa*, 962 P.2d 909, 915 (Ariz. 1998) (defendant's liability is not "limited by apportioning damages, but only by apportioning fault"); *Quynn v. Hulsey*, 850 S.E.2d 725, 729 (Ga. 2020) (jury apportions damages between multiple tortfeasors "according to the percentage of their fault"); *Fed. Deposit Ins. Corp. v. Loudermilk*, 826 S.E.2d 116, 126 (Ga. 2019) (same); *Owens Corning Fiberglas Corp. v. Parrish*, 58 S.W.3d 467, 479 (Ky. 2001) ("If . . . the evidence does not permit apportionment of the damage between separate causes, then comparative fault principles apply, and the trial court should instruct the jury to apportion damages according to the proportionate fault of the parties."); *Carter v. Wallace & Gale Asbestos Settlement Tr.*, 96 A.3d 147, 160 (Md. Ct. App. 2014) (noting that where injury is divisible and apportionment is a fact question, burden rests with defendant to show how damages should be apportioned); *Campione v. Soden*, 695 A.2d 1364, 1370-71 (N.J. 1997) (discussing apportionment and reasoning that the question of apportionment is for the jury); *Pratt v. Amisub of SC, Inc.*, 912 S.E.2d 268, 282-83 (S.C. Ct. App. 2025) (explaining South Carolina's joint tortfeasor statutory scheme, including role of jury in apportioning fault and damages); *see also* Restatement (Second) of Torts § 433A (explaining difference between apportionment for divisible injuries with an evidentiary basis for determining contribution, versus indivisible harms where evidence that guides apportionment of damages is lacking).

Defendants separately criticize the Districts for relying on interrogatory responses. *E.g.,* Defs.' Tucson Mot. at 33, 35. These criticisms are unavailing. Experts are permitted to rely upon sworn interrogatory responses, and Rule 56 expressly contemplates that interrogatories may be used to demonstrate a genuine dispute of material fact that will defeat a motion for summary judgment. *See*

Correct Transmission, LLC v. Nokia of Am. Corp., 2024 WL 1289821, at *5 (E.D. Tex. Mar. 26, 2024); Fed. R. Civ. P. 56(c)(1)(A). If Defendants wish to challenge the sufficiency of Plaintiffs' interrogatory responses, which have been verified under penalty of perjury, they can present that argument to the jury.

2. Defendants can be required to compensate Districts for the diversion of school resources caused by their misconduct.

Defendants' challenge to the Districts' diversion-of-resources damages fails for multiple, independently sufficient reasons. Across their six motions for summary judgment, Defendants fundamentally misstate the law in asserting that the Districts cannot recover the costs of employee time diverted to address the disruptions to instructional time and school operations caused by Defendants' social media platforms. Defendants' position rests on inapposite authority and a mischaracterization of compensable damages, improperly suggesting that the Districts may recover only new or additional expenditures incurred as a direct result of Defendants' misconduct. The law is clear that diversion of resources itself constitutes a recognized and compensable injury, and courts have long permitted recovery of such damages. *See Convoy Co. v. Sperry Rand Corp.*, 672 F.2d 781, 785 (9th Cir. 1982).

As an initial matter, the Ninth Circuit has squarely rejected Defendants' improper contention that the Districts cannot recover for employee time lost to addressing harms caused by Defendants, simply because the Districts would have paid those employees' salaries and benefits regardless. In *Convoy Co. v. Sperry Rand Corp.*, the plaintiff sought damages for "the hours its salaried personnel spent supervising [a malfunctioning] computer system." *Id.* at 783. On appeal, the defendant argued that such damages were impermissible because "salaried supervisory staff costs . . . [are] not allowable as a matter of law" and the plaintiff "would have paid the staff's salary in any event." *Id.* at 785. The Ninth Circuit expressly rejected that argument, holding that "the issue is not whether [the plaintiff] would have paid the supervisors' salaries if the defendant had not breached the contract, but whether the breach deprived Convoy of the services it paid for." *Id*.

Indeed, "[m]ost courts that have considered the issue have concluded that a plaintiff can recover the value of employees' lost services as damages in a contract or tort action, ***even when it had not shown that it incurred additional expenses or lost profits.***" *State v. Rouse*, 647 N.W.2d 286, 289 (Wis. Ct. App. 2002) (emphasis added) (collecting cases); *see also Zayo Grp., LLC v. Sols. Fiber Optic, Inc.*, 2025 WL

2157902, at *4 (E.D. Va. July 11, 2025), *report and recommendation adopted*, 2025 WL 2155792 (E.D. Va. July 29, 2025) (ordering damages including "internal labor costs"); *Mobile Conversions, Inc. v. Allegheny Ford Truck Sales*, 2014 WL 7369898, at *6 (W.D. Pa. Dec. 29, 2014); *Dunn Appraisal Co. v. Honeywell Info. Sys. Inc.*, 687 F.2d 877, 883 (6th Cir. 1982) (upholding award of 1/5 of a CEO's salary based on time spent addressing fallout from defendants' conduct).

Specifically, Kentucky courts recognize that lost time and opportunity costs incurred while mitigating the consequences of another's wrongful conduct constitute a cognizable form of damages in negligence actions. *See Lurry v. PharMerica Corp.*, 2024 WL 2965642, at 3 (W.D. Ky. June 12, 2024) ("lost time and opportunity costs associated with attempting to mitigate the actual consequences of the defendant's conduct" are recoverable damages); *see also Shulz v. Chadwell*, 558 S.W.2d 183, 187 (Ky. Ct. App. 1977) (holding that where one partner became unavailable for work, the amount paid to the other partner to perform the same work was "sufficient evidence of probative value on the issue of lost time"). Defendants' cited authorities do not support a contrary conclusion. *See Schwartz v. Hasty*, 175 S.W.3d 621, 625 (Ky. Ct. App. 2005) (not addressing lost employee time); *Gassaway Constr. Co. v. Gentry*, 264 S.W.2d 658, 659 (Ky. 1954) (finding insufficient evidence to establish loss-of-time damages).

Similarly, Georgia courts permit plaintiffs to seek recovery of portions of employee salaries, such as those sought here, and reject arguments that permitting such damages would allow recovery of costs the plaintiff would have otherwise incurred. *See Aaron's, Inc. v. MDC Grp., Inc.*, 2010 WL 11505919, at *7 (N.D. Ga. Dec. 15, 2010). Similarly, the Third Circuit, applying New Jersey law, upheld an award for compensatory damages including "the value of the time [Plaintiff's] employees expended in attempting to reduce or avert the damage flowing from the" defendant's conduct. *Comdyne I, Inc. v. Corbin*, 908 F.2d 1142, 1146, 1150 (3d Cir. 1990). Courts in Arizona have also upheld damages seeking compensation for lost employee time. *See PivotHealth Holdings LLC v. Horton*, 2025 WL 1865788, at *2 (D. Ariz. July 7, 2025) (finding standing based on damages including "employee time" that was "incurred in responding to" the defendant's conduct); *IceMOS Tech. Corp. v. Omron Corp.*, 2020 WL 1083817, at *4 (D. Ariz. Mar. 6, 2020) (permitting plaintiff to present evidence at trial "seeking costs of the salaries of its employees for time spent working" to address defendant's conduct).

Defendants' own cited authority confirms that lost employee time is a compensable form of

damages under South Carolina law. *See Charleston Lumber Co. v. Miller Hous. Corp.*, 458 S.E.2d 431, 437 (S.C. Ct. App. 1995) ("The Millers alleged damages of lost employee time based on their personnel having to spend numerous hours each month checking and correcting the bids versus the actual charges. We find that employee time is a compensable damage."). While Defendants suggest -- without citing any authority—that *Charleston Lumber*'s holding is limited to fraud claims, courts routinely recognize lost-time damages in negligence actions. *See In re Blackbaud, Inc., Customer Data Breach Litig.*, 567 F. Supp. 3d 667, 687 (D.S.C. 2021) (finding plaintiffs alleged cognizable injury under South Carolina law for time and money spent mitigating injuries arising from defendant's negligence).

Finally, courts applying Maryland law have expressly rejected the same argument Defendants advance here—that no loss is incurred because an employer would have paid the employees anyway. *See Under Armour, Inc. v. Ziger/Snead, LLP*, 158 A.3d 1134, 1138 (Md. Ct. Spec. App. 2017) (holding that "diverted employee time [is] a compensable loss"). As Maryland courts have explained, because "diverted employee time" prevents employees from rendering "other service to the [plaintiff]," such costs are recoverable. *MacDonald v. Patriot, LLC*, 2017 WL 1788115, at *12 (Md. Ct. Spec. App. May 5, 2017); *see also KeraLink Int'l, Inc. v. Stradis Healthcare, LLC*, 2021 WL 5832242, at *1 (D. Md. Nov. 2, 2021), *aff'd sub nom. KeraLink Int'l, Inc. v. Geri-Care Pharms. Corp.*, 60 F.4th 175 (4th Cir. 2023) (entering judgment for damages including "costs of employee time").

Defendants rely on authority disallowing recovery for lost *personal* time. That authority is inapposite: Those cases involved plaintiffs seeking compensation for their own inconvenience, not for the loss of employee services suffered by an organizational plaintiff. *See Quinalty v. FocusIT LLC*, 2024 WL 342454, at *4 (D. Ariz. Jan. 30, 2024) (plaintiff alleged lost personal time responding to data breach); *Griffey v. Magellan Health Inc.*, 562 F. Supp. 3d 34, 45 (D. Ariz. 2021) (same); *In re Gen. Motors LLC Ignition Switch Litig.*, 339 F. Supp. 3d 262 (S.D.N.Y. 2018) (plaintiffs alleged lost personal time seeking repair of their vehicle). The Districts' claim is different. The Districts do not seek recovery for personal inconvenience, but for the value of employee time and services diverted from their intended educational and operational functions to address the disruptions caused by Defendants' platforms. As the Sixth Circuit recognized, "[employer] was paying [employee], and the time he spent was [employer]'s time, not his own." *Dunn Appraisal Co. v. Honeywell Info. Sys. Inc.*, 687 F.2d 877, 884 (6th Cir. 1982).

As Dr. Ward explained, his calculation of damages based on diverted employee time "is based on the fundamental economic concept of opportunity cost," and the harm to the Districts is "the value of the foregone alternative." *E.g.*, Ex. 1083 ¶ 8. To quantify the cost of these foregone activities, "[e]conomists often estimate the opportunity cost of time using the individual's wage." *Id.* ¶ 9. This is a widely-accepted method of calculating the opportunity costs in an educational setting. As Dr. Ward explained, where an intervention "requires teachers to reallocate their time from instruction, that time should be counted as an opportunity cost to the school represented by the teachers' wages and benefits for that time." Ex. 1084 ¶ 17 (quoting Levin et al., *Economic Evaluation in Education: Cost-Effectiveness and Benefit-Cost Analysis*, 4th ed. 2017, at 69).

In short, the diversion of teacher and staff time from their educational mission, to address the disruptions caused by Defendants' social media platforms, constitutes a compensable form of harm, recognized both by established precedent and economic principles. While Defendants fault the Districts for responding to this novel and externally imposed harm with the resources available to them, state and federal law uniformly recognize that when a defendant's conduct deprives a plaintiff of the services of its personnel, that deprivation is a cognizable form of damages. *See Convoy Co.*, 672 F.2d at 785. Defendants' motions should be denied to the extent they seek to preclude the Districts' recovery for the substantial loss of employee time and resources diverted to address the harms caused by Defendants' platforms.

3. Mr. Klein's survey raises disputed issues of fact regarding the Districts' lost time damages.

Defendants incorrectly contend that the Districts lack evidence of their lost time damages by challenging the teacher survey conducted by Mr. Klein. As explained in Plaintiffs' Omnibus Daubert Opposition, Mr. Klein—an MIT-educated expert in market research with over fifty years of experience and more than 1,000 surveys administered—designed and executed the survey in accordance with well-accepted principles of market research. Its findings provide reliable, admissible evidence supporting the Districts' damages and raise disputed issues of fact regarding lost instructional time. *See* SD Daubert Opp. § III.B; Ex. 1085 ¶ 41 (Breathitt: scheduled classroom time diverted 1.4-5.2% for middle school, 5.2-3.9% for high school); Ex. 1086 ¶ 41 (Charleston: scheduled classroom time diverted 5.5-7.1% for middle school, 4.1-4.3% for high school); Ex. 1087 ¶ 41 (DeKalb: scheduled classroom time diverted 1.6-5.2%

for middle school, 4.6-11.7% for high school); Ex. 1088 ¶ 41 (Harford: scheduled classroom time diverted 1.9-5.7% for middle school, 3.1-10.6% for high school); Ex. 1089 ¶ 41 (Irvington: scheduled classroom time diverted 4.7-2.2% for middle school, 3.7-9.9% for high school); Ex. 1090 ¶ 41 (Tucson: scheduled classroom time diverted 2.6-5.2% for middle school, 3.7-14.8% for high school).

Mr. Klein's survey quantifies the instructional time lost because of Defendants' conduct based on the accounts of those with first-hand knowledge—teachers and administrators directly impacted by Defendants' platforms. While Defendants seek to dismiss these educator accounts, Mr. Klein's analysis offers a scientifically valid and methodologically sound measure of the Districts' diverted resources and lost time. *See* SD Daubert Opp. § III.B. Courts routinely recognize that survey evidence based on employee self-reporting can provide a reasonable approximation of damages where detailed contemporaneous records are unavailable. *See, e.g.*, *Alcantar v. Hobart Serv.*, 2013 WL 156530, at 4 (C.D. Cal. Jan. 15, 2013) ("It would be unreasonable to expect employees to maintain records of every time they were provided a meal break," and "survey evidence … is relevant to give an approximation of damages.").

Ultimately, Defendants' criticisms of the Klein survey go to weight, not admissibility. The survey easily satisfies Rule 702 and creates genuine issues of material fact sufficient to preclude summary judgment on the issue of lost time. *See id.*; *Fortune Dynamic, Inc. v. Victoria's Secret Stores Brand Mgmt., Inc.*, 618 F.3d 1025, 1037 (9th Cir. 2010) (survey admissible under Rule 702 raised triable issue of fact).

D. The Districts' claims for future damages are legally cognizable.

1. The Districts proffer relevant evidence supporting a claim for future damages.

In each bellwether jurisdiction, a plaintiff may obtain an award of future damages if it proves ongoing injury and the costs of remediation. *See, e.g.*, *May v. Holzknecht*, 320 S.W.3d 123, 128 (Ky. Ct. App. 2010); *Haltiwanger v. Barr*, 186 S.E.2d 819, 820 (S.C. 1972); *Wilder v. Blue Ribbon Taxicab Corp.*, 719 S.E.2d 703, 708 (S.C. Ct. App. 2011); *Zakas v. Jackson*, 835 S.E.2d 371, 373 (Ga. Ct. App. 2019); *DiLeo v. Nugent*, 592 A.2d 1126, 1134–35 (Md. Ct. Spec. App. 1991); *Colon v. Robinson*, 2014 WL 7466566 at *6 (N.J. Super. Ct. App. Div. Jan. 6, 2015); *Saide v. Stanton,* 659 P.2d 35, 38 (Ariz. 1983). Here, the Districts properly seek future damages under their negligence claims by presenting evidence that (1) Defendants' platforms are reasonably certain to cause ongoing injury and (2) the costs of

addressing those harms. *See, e.g.*, Ex. 1000; *see also* Ex. 1098. That is all the law requires.

The Districts have presented expert testimony from Dr. Hoover explaining that the Defendants have caused numerous ongoing harms that are reasonably certain to recur. *See, e.g.*, Ex. 1000 ¶¶ 54–55 (describing "frequent and pervasive" social-media interruptions that undermine instruction), ¶ 61 (explaining the "growing burden on teachers" to manage behavioral fallout), ¶ 77 (describing diversion of "limited funding and staffing resources" to address social-media impacts); Ex. 1096 ¶ 55 (opining that Charleston County requires additional resources to address social-media harms); *accord*; Ex. 1091 ¶ 65; Ex. 1092 ¶ 55, Ex. 1093 ¶ 56; Ex. 1094 ¶ 59; Ex. 1095 ¶ 66; *see also* Ex. 1000 ¶¶ 30–84.

Dr. Hoover next opines that her detailed, 15-year strategic plans are necessary to prevent these future injuries. *See* Ex. 1091 ¶¶ 106, 110; Ex. 1093 ¶¶ 97, 101; Ex. 1094 ¶¶ 100, 104; Ex. 1095 ¶¶107, 111Dr. Leslie then quantifies the costs necessary for each District to implement those plans. *See* Ex. 609, Ex. 610, Ex. 611, Ex. 612, Ex. 613, Ex. 614. This evidence alone establishes a genuine issue of material fact on future damages. *See Haltiwanger*, 186 S.E.2d at 820.

Defendants cite no authority requiring plaintiffs to prove they can fund mitigation measures absent a damages award. No such rule exists, and for good reason: It would be illogical for Defendants to be absolved just because the entities they harm are underfunded public institutions. Defendants cite three groups of cases, each less relevant than the last.

First, they cite cases in which damages were denied because the asserted future injury was speculative, not reasonably certain. *See Mauro v. Raymark Indus., Inc.*, 561 A.2d 257 (N.J. 1989); *Battaglia v. United Parcel Serv., Inc.*, 70 A.3d 602 (N.J. 2013); *J.N. Legacy Grp., Inc. v. City of Dallas*, 745 S.E.2d 721, 729 (Ga. Ct. App. 2013); *Pierce v. Johns-Manville Sales Corp.*, 464 A.2d 1020, 1026 (Md. 1983); *Davidson v. Miller*, 344 A.2d 422, 427 (Md. 1975). These cases are inapposite, given the Districts *have* presented evidence indicating their future injuries are reasonably certain. *See, e.g.*, Ex. 1000 ¶ 8 ("student social media use has had—and continues to have—a profound and detrimental impact"); Ex. 1099 ¶¶ 72, 135, 175. At a minimum, whether a future injury is reasonably certain is a question for the jury. *See Mauro*, 116 N.J. at 139.

Second, pertinent to Tucson's case, Defendants cite Arizona cases involving lost profits—*Gilmore v. Cohen*, 386 P.2d 81 (Ariz. 1963); *McAlister v. Loeb & Loeb, LLP*, 571 P.3d 891 (Ariz. 2025); *Rancho*

Pescado, Inc. v. Nw. Mut. Life Ins. Co., 680 P.2d 1235 (Ariz. Ct. App. 1984); and *McNutt Oil & Refin. Co. v. D'Ascoli*, 281 P.2d 966, 970-971 (Ariz. 1955). But as those cases explain, a plaintiff needs "such precision as, from the nature of the claim and available evidence, is possible." *Gilmore*, 386 P.2d at 83. And lost profits are distinct because they are amenable to "mathematical precision"—they may be proven through "books of account," "informal memoranda of previous transactions," or "past income tax returns." *Gilmore*, 386 P.2d at 82–83. The Districts do not seek lost profits and, in any event, Dr. Leslie has provided precise calculations of the costs necessary to address the ongoing harms caused by Defendants' platforms. *See* Ex. 614 ¶¶ 19-20.

Third, Defendants cite cases that simply do not involve future damages at all, and hence are entirely irrelevant to the question at hand. *See Nobles v. Jiffy Mkt. Food Store Corp.*, 579 S.E.2d 63 (Ga. Ct. App. 2003) (concerning liquidated damages award, not future losses); *see also J&W Corp. v. Broad Creek Marina, LLC*, 896 S.E.2d 328 (S.C. Ct. App. 2023); *Accessory Overhaul Grp., Inc. v. Mesa Airlines, Inc.*, 994 F. Supp. 2d 1296 (N.D. Ga. 2014); *Carter v. Wallace & Gale Asbestos Settlement Tr.*, 96 A.3d 147, 157 (2014); *Coury Bros. Ranches, Inc. v. Ellsworth*, 446 P.2d 458 (Ariz. 1968); *Hirsh v. Manley*, 300 P.2d 588 (Ariz. 1956).

Here, the evidence establishes ongoing harm with reasonable certainty. Dr. Hoover's plans were specifically designed to address the continuing, systemic injuries Defendants' platforms cause to school operations, learning, and student well-being. *See* Ex. 1000 ¶ 5. Specifically, Dr. Hoover determined that Defendants' platforms continue to damage the Districts by negatively affecting:

- **School operations** ("reshaped the educational landscape in ways that extend well beyond individual student behavior," Ex. 1000 ¶ 83);
- **School climate and environment** ("pervasive negative impact on schools and the school environment," *id.* ¶ 34);
- **Learning and performance** ("many teachers report that social media use interferes with their teaching and students' learning," *id.* ¶ 45; "growing frustration with managing screen use in class," *id.* ¶¶ 60, 80, 82);
- **Teaching effectiveness and classroom dynamics** ("the strain Defendants' platforms place on teacher-student relationships," *id.* ¶ 53);
- **Teacher morale and job satisfaction** ("teachers often experience a decline in morale when faced with the constant distractions caused by students' social media use," *id.* ¶ 58); and

- **Student mental health** ("[s]chools are increasingly burdened with managing the mental health consequences of social media overuse," *id*. ¶ 78; *see id*. ¶¶30–82; Exs. 1091–1096 ¶ 4 (each describing how districts have been forced to "expend and redirect already limited resources" to address these harms)).

Dr. Hoover concludes these harms will persist and intensify absent substantial intervention, warning that without implementation of her plans, the Districts will "perpetuat[e] ineffective approaches" and remain unable "to coordinate services, evaluate impact, and make data-informed decisions." *See* Ex. 1000 ¶¶ 8–9, 125; *see also* Ex. 1102 ¶¶ 35, 46, 60; Ex. 1103 ¶¶ 32, 51; Ex. 1101 ¶¶ 33, 58; Ex. 1104 ¶ 44; Ex. 1105 ¶¶ 39, 137; Ex. 1100 ¶¶ 56, 129. She emphasizes that her 15-year, multi-tiered framework is necessary to prevent and mitigate escalating harms within school environments. Combined with Dr. Leslie's detailed cost analysis, this evidence provides a concrete, non-speculative basis for a future-damages award.

Defendants' assertion that Dr. Hoover's plans are merely optional "recommendations" mischaracterizes her work (and diminishes what is, at a minimum, a genuine dispute of fact for the jury to resolve). Dr. Hoover makes clear that the measures she proposes are evidence-based requirements, not flexible suggestions. As she explains, "the staffing and cost recommendations [she] present[s] are not aspirational targets or flexible guidelines; they represent the level of investment and capacity required to adequately address the harms due to social media use in school communities." Ex. 1101 ¶ 41. Across her reports, Dr. Hoover underscores that these recommendations are grounded in empirical data, implementation science, and public-health precedent, reflecting the actual scale of intervention needed to remediate the harms caused by Defendants' platforms. *See also* Ex. 1100 ¶ 189 ("not aspirational" but "grounded in evidence of what is needed to meet the behavioral and mental-health needs of students"); Ex. 1102 ¶ 45 (providing "evidence-informed recommendations for the staffing, training, and system design needed to meet" the harms); Ex. 1105 ¶ 32 (recommendations are "realistic and proportionate to respond" to social-media harms); Ex. 1103 ¶ 41; Ex. 1104 ¶ 32.

Nor is it surprising that the Districts have not yet implemented Dr. Hoover's plans. These are resource-constrained public school systems seeking damages precisely because they cannot fund such measures without relief. Dr. Hoover's role was not to design a plan limited by current budgets but to identify what is necessary to remediate the harms Defendants have caused. *See, e.g.*, Ex. 1105 ¶ 27 ("My

role is not to budget based on current constraints, but to provide an evidence-informed benchmark for the staffing levels necessary to meet the scale and specificity of social media-related student distress."). That other districts have not yet implemented identical frameworks is immaterial (and, in any event, may just demonstrate their comparable budget constraints); Dr. Hoover's plans are tailored to these Plaintiffs' unique needs and the specific harms caused by Defendants' platforms. The fact that no other district has had the funding or occasion to request such plans underscores their necessity, not their novelty. In any event, Defendants' argument ignores the over one thousand other school districts who, like the Districts, have brought suit against Defendants seeking comparable relief.

Defendants' claim that Districts cannot hire sufficient personnel is likewise baseless. Dr. Hoover directly refutes it: "[C]ontrary to claims that the proposed staffing levels are unattainable, current data reflect a steady flow of qualified school mental-health professionals entering the field." Ex. 1099 ¶ 48; *see also* Ex. 1102 ¶ 41; *accord* Exs. 1100 ¶ 40; 1101 ¶ 38; 1103 ¶ 34; 1104 ¶ 34; 1105 ¶ 31.

Finally, Defendants' argument that Tucson and Charleston have not yet budgeted for social-media-specific interventions is irrelevant. As Dr. Hoover explains, these Districts "lack the resources and infrastructure to address the modern, external threat to student well-being and school functioning" within their standard budgets. Ex. 1100 ¶ 175; *see also id*. ¶ 182 ("that social-media harms are not prominently featured in older internal documents or district budgets further supports my conclusion that districts have not had the systems, staffing, or awareness to adequately monitor or address these harms"); Ex. 1101 ¶ 128 ("the fact that a district has not fully incorporated a new and emerging harm into its formal systems ... does not mean the harm does not exist or is not substantial"). Similarly, Irvington's historic understaffing has only been exacerbated by Defendants' misconduct. *See* Ex. 1105 ¶¶ 26–28.

In sum, the Districts have presented robust, expert-supported evidence of ongoing, reasonably certain future harm, and the precise costs necessary to remedy them. These are issues of fact for a jury, not matters for summary judgment. *Fuller v. Idaho Dep't of Corr.*, 865 F.3d 1154, 1161 (9th Cir. 2017).

2. Breathitt, DeKalb, Harford, and Tucson's nuisance claims permit the abatement they seek.

a) Abatement is a distinct equitable remedy designed to eliminate ongoing and future public harms.

Breathitt, DeKalb, Harford, and Tucson seek the equitable remedy of abatement to address ongoing and future harms caused by Defendants' nuisance. *See* Restatement (Second) of Torts § 821C(2) (1979). In addition to asserting negligence claims, Breathitt, DeKalb, Harford, and Tucson bring claims for public nuisance, which rest on a different legal theory. Negligence requires proof that Defendants owed and breached a duty. *See* MTD Order, 756 F. Supp. 3d at 972. Remedies are limited to money damages (any may include future damages, *see supra*). Nuisance requires proof that Defendants created "an unreasonable interference with a right common to the general public," and, if proven, authorizes the Court to fashion an equitable remedy. *See* Restatement (Second) of Torts §§ 821B–C (1979).

Abatement is a court-crafted equitable remedy, wholly distinct from damages. *City of Huntington, W. Virginia v. AmerisourceBergen Drug Corp.*, 2025 WL 3009526, at *12 (4th Cir. Oct. 28, 2025) ("[P]ublic nuisance claims serve [the] function [of] focusing on harm suffered by the public more generally and seeking abatement of that harm through the court's equitable authority."); *State of Wash. v. McKesson Corp. et al.*, No. 19-2-06975-9 (Sup. Ct. Wash. July 6, 2021) (granting motion to strike jury demand for abatement because it is "equitable in nature, and should be tried to the Court"). As explained in *In re Nat'l Prescription Opiate Litig.*, "[t]he goal [of abatement] is not to compensate the harmed party for harms already caused by the nuisance. This would be an award of damages. Instead, an abatement remedy is intended to compensate the plaintiff for the costs of rectifying the nuisance, going forward." 2019 WL 4043938, at *1 (N.D. Ohio Aug. 26, 2019). Because abatement may "be more certain, prompt, or efficient than the monetary damages [they] seek[], but may ultimately not attain," it is proper for the Districts to pursue both damages and abatement. *See Coleman v. Mondelez Int'l Inc.*, 554 F. Supp. 3d 1055, 1065 (C.D. Cal. 2021) (plaintiffs may pursue equitable relief and damages under separate claims).

Here, the Districts ask the Court to create an abatement plan to eliminate the nuisance in their schools. Guided by the testimony of Dr. Hoover and Dr. Leslie, this plan would be a post-trial, Court-crafted remedy akin to an injunction. *See id*. Unlike a damages award, abatement allows the Court to order and oversee implementation to address Defendants' ongoing harms. *See Andino v. Apple, Inc.*, 2021 WL

1549667, at *5 (E.D. Cal. Apr. 20, 2021) ("Money damages are an inadequate remedy for future harm.").

Defendants argue that potential recovery of future damages bars abatement. They rely on *Sonner v. Premier Nutrition Corp.* but misread its holding. *Sonner* merely provides that a plaintiff cannot "secure equitable" relief ***if*** an adequate remedy at law exists. 971 F.3d 834, 844 (9th Cir. 2020) (emphasis added). At trial, a plaintiff may pursue both (as cases interpreting *Sonner* make clear). *Cepelak v. HP Inc.*, 2021 WL 5298022, at *2 (N.D. Cal. Nov. 15, 2021) (allowing damages and equitable relief in the alternative); *Coleman*, 554 F. Supp. 3d at 1065; *Ostrovskaya v. St. John Knits, Inc.*, 2022 WL 2102895, at *5 (C.D. Cal. Mar. 31, 2022); *Roper v. Big Heart Pet Brands, Inc.*, 510 F. Supp. 3d 903, 917 (E.D. Cal. 2020). ***If*** a jury ultimately awards future damages duplicative of the abatement sought, the Court may decline to impose an additional equitable remedy. In various cases cited by Defendants, that was the case: The equitable relief sought was wholly duplicative of the damages requested. *See Wood v. Marathon Refining Logistics Serv. LLC*, 2024 WL 2242688, at *10 (N.D. Cal. Mar. 21, 2024) (barring plaintiff from seeking equitable restitution that merely duplicated past damages); *Rhynes v. Stryker Corp.*, 2011 WL 2149095, at *4 (N.D. Cal. May 31, 2011) (plaintiff failed to show damages would not provide full compensation).[17] By contrast, the abatement the Districts seek uniquely addresses ongoing and future harms by allowing the Court to fashion a remedy that eliminates the nuisance.

In any event, the determination of whether abatement is or is not duplicative of a damages award is properly made post-trial. *See Teutscher v. Woodson*, 835 F.3d 936, 957 (9th Cir. 2016) (holding that a plaintiff may seek backward-looking damages from a jury and a "forward-looking remedy . . . from the court in equity"). Accordingly, Breathitt, DeKalb, Harford, and Tucson may pursue equitable abatement through their public nuisance claims.

b) Dr. Hoover's strategic plans provide the framework for crafting an abatement remedy.

As explained in the Districts' opposition to Defendants' motion seeking to exclude Dr. Hoover and other experts ("SD Daubert Opp."), Dr. Hoover's plans were developed consistently with her standard

[17] Defendants also cite two entirely inapposite cases in support of their argument—*Robinson v. C.R. Bard, Inc.*, 2016 WL 3361825 (N.D. Cal. June 17, 2016) (where the plaintiff had not pursued any timely claims) and *Mitsubishi Int'l Corp. v. Cardinal Textile Sales, Inc.*, 14 F.3d 1507, 1518 (11th Cir. 1994) (concerning the suis generis procedural and equitable context of a pre-judgment constructive trust on assets).

practice and are grounded in public health and implementation science. They outline in specific detail the staffing and programs needed to prevent and mitigate the harms caused by Defendants—providing comprehensive, district-wide interventions, including universal policies for all students, staff, and families, and more intensive programs for those with greater needs. This approach is essential because the harms from Defendants' platforms "pervade the school[s]" and require proactive services "for those students who may not have yet used a platform directly themselves." Ex. 1097B at 623:10–624:4; *see also* Ex. 1000 ¶ 34. Dr. Hoover further opines that a 15-year timeline is the "necessary duration for implementing a comprehensive, district-led strategic plan to prevent and mitigate the negative impact of social media on school districts, schools, the school environment, student well-being and learning," given comparable public health initiatives, the developmental K–12 timeframe, and system change requirements. Ex. 1000 ¶ 7; *see also* ¶ 33 (describing a 15-year integrated mitigation plan).

Dr. Hoover's strategic plans provide the Court with detailed guidance for crafting an abatement remedy for Breathitt, DeKalb, Harford, and Tucson. *See* Ex. 1091 ¶¶ 106, 110; Ex. 1093 ¶¶ 97, 101; Ex. 1094 ¶¶ 100, 104; Ex. 1095 ¶¶ 107, 111; *see also In re Nat'l Prescription Opiate Litig.*, 2019 WL 4043938, at *1 (N.D. Ohio Aug. 26, 2019) (noting that abatement is an equitable remedy fashioned by the Court using expert testimony and trial evidence). Further, Dr. Leslie calculated the cost to implement Dr. Hoover's plan for each District. This evidence should also assist the Court in crafting an abatement remedy, as monetary funds are an appropriate form of abatement. *See infra*. In short, the Districts have offered precisely what is required—expert guidance to assist the Court in crafting an abatement remedy

(1) The abatement remedy may include monetary relief.

Defendants wrongly argue that abatement cannot include monetary relief. Courts have long rejected the "myth" that equitable remedies never involve payment: "the assumption that 'equitable remedies are always orders to act or not to act, rather than to pay, is a myth; equity often orders payment.'" *United States v. Apex Oil Co.*, 579 F.3d 734, 736 (7th Cir. 2009).

As an equitable remedy, abatement may be "mold[ed] to the circumstances of [the] particular case." *Weinberger v. Romero-Barcelo*, 456 U.S. 305, 312 (1982). Courts addressing public health crises have recognized that abatement may include a monetary fund. In the opioid litigation, the Fourth Circuit held that "the remedy of abatement . . . permit[s] a monetary award to fund . . . abatement efforts,"

emphasizing that "[c]ourts have broad powers to effect equitable relief." *City of Huntington, W. Virginia v. AmerisourceBergen Drug Corp.*, 2025 WL 3009526, at *20 (4th Cir. Oct. 28, 2025). Likewise, in *In re JUUL Labs, Inc., Mktg., Sales Pracs., & Prods. Liab. Litig.*, the court found that a nuisance abatement remedy may "require defendants to expend the money necessary to abate the nuisance" and ordered contributions to an abatement fund to address youth vaping. 497 F. Supp. 3d 552, 653 (N.D. Cal. 2020). That is precisely what the Districts seek here—a post-trial, Court-crafted remedy guided by Dr. Hoover's and Dr. Leslie's testimony regarding the measures and costs necessary to abate the nuisance. Defendants' effort to distinguish *JUUL* fails; while the *JUUL* plaintiffs proposed only an outline, the Districts have presented detailed, expert-supported abatement plans for the Court's consideration.

Recent decisions confirm the same principle. In *Mayor and City of Baltimore v. Purdue Pharma, L.P.*, No. 24-C-18-000515, slip op. at 1, 23 (Md. Cir. Ct. Aug. 8, 2025), the court crafted a post-trial abatement plan funded by monetary payments to hire staff and expand treatment and harm-reduction services. Similarly, Dr. Hoover's plans propose staffing and training programs to prevent and mitigate harms caused by Defendants' platforms. *See, e.g.*, Ex. 1000 ¶¶ 6, 83–84, 95–122; Ex. 1093 ¶¶ 74–85. Defendants attempt to distinguish the *Purdue Pharma* plan by noting that the court did not fund *every* proposed treatment and service cost. This misses the point. The *Purdue Pharma* court conducted a fact-based, post-trial inquiry, guided by the jury's findings, to determine which treatments and services should be included—and it did incorporate multiple treatment and recovery programs into its plan. This Court should engage in a similar post-trial evaluation to determine the appropriate scope of the abatement plan.

In re Nat'l Prescription Opiate Litig., 622 F. Supp. 3d 584, 606 (N.D. Ohio 2022), *rev'd on other grounds*, 2025 WL 354758 (6th Cir. Jan. 31, 2025), likewise held that abatement may require payment, explaining that "eliminating a hazard that continues to cause prospective harm . . . will, in virtually all cases, cost a liable defendant some amount of money," and that an equitable abatement award "forces a liable defendant to clean up the mess it made." *Id.* Again, that is exactly what the Districts seek. While Defendants disingenuously assert that *In re Nat'l Prescription Opiate Litig.* was vacated, no appellate court has overturned any of the court's rulings on abatement. The appeal addressed only whether an Ohio state statute barred the underlying public nuisance claim, an issue wholly irrelevant here.

Other courts agree that monetary abatement remains equitable and confirm that this Court may—and should—fashion a monetary abatement remedy tailored to these harms. See *State of Wash. v. McKesson Corp.*, No. 19-2-06975-9, slip op. at 2 (Wash. Super. Ct. July 6, 2021) (quoting *State ex rel. Dep't of Ecology v. Anderson*, 620 P.2d 76 (Wash. 1980)) ("a claim remains equitable even though a money judgment might form a part of the relief asked"); *People v. ConAgra Grocery Prods. Co.*, 227 Cal. Rptr. 3d 499 (Ct. App. 2017) (requiring defendants to fund an abatement account to carry out remediation).

(2) Defendants' state law authorities are inapposite.

Defendants' reliance on state cases to limit this Court's authority is misplaced. Their own authority, *Sonner*, confirms that "state law cannot expand or limit a federal court's equitable authority," even in diversity cases. 971 F.3d at 841. Federal courts possess "broad powers to tailor and fashion appropriate equitable remedies." *United States v. Price*, 688 F.2d 204, 211 (3d Cir. 1982); *Guthrie v. Transamerica Life Ins. Co.*, 561 F. Supp. 3d 869, 873 (N.D. Cal. 2021).

Even if state law were relevant, state courts have recognized that abatement can extend beyond simply stopping or removing the nuisance. For example, *Brown v. City of Phoenix*, 557 P.3d 321, 327 (Ariz. Ct. App. 2024) distinguishes between the removal and the abatement of nuisances and dealt with an injunction requiring both the removal of certain individuals and the cleanup of hazardous materials they left behind. *City of Safford v. Seale*, 2009 WL 3390172, at *2 (Ariz. Ct. App. Oct. 21, 2009) also recognizes that abatement can include remediation as well as removal. Similarly, while *Spaw, LLC v. City of Annapolis*, 156 A.3d 906, 931 (Md. Ct. App. 2017) did not consider the proper scope of abatement remedies, it recognized in dicta that requiring the defendant to not just halt the offending conduct but affirmatively undertake action to undo the effects of the violation would be an appropriate form of abatement. *See also Adams v. Comm'rs of Town of Trappe*, 102 A.2d 830 (Md. Ct. App. 1954) (requiring defendant to not just stop sale of gasoline but remove tank and pump).

Additionally, none of the cases cited by Defendants held that enjoining conduct is the only acceptable form of abatement. In *Superior Farm Mgmt., L.L.C. v. Montgomery*, 513 S.E.2d 215, 217-18 (Ga. 1999), the nuisance was planned but had not yet begun, so no other form of abatement other than an injunction was possible. *See also Brandes v. Mitterling*, 196 P.2d 464 (Ariz. 1948) (plaintiffs sought only injunction against future use); *Cactus Corp. v. State ex rel. Murphy*, 480 P.2d 375 (Ariz. Ct. App. 1971)

(same, and nuisance was present for less than two weeks). Meanwhile, *Martin v. Howard Cnty.*, 709 A.2d 125 (Md. 1998) and *J.D. Jewell, Inc. v. Hancock*, 175 S.E.2d 847 (Ga. 1970), address, respectively, the difference between legal and equitable remedies and jurisdiction and venue, not the appropriate forms of abatement. Likewise, *Green Meadows Hous. Partners, LP v. Macon-Bibb Cnty.*, 906 S.E.2d 430, 441 (2024) turned on the statutory scope of receiverships, not the breadth of available abatement remedies.

(3) Defendants' attacks on Dr. Hoover's plans lack merit.

Defendants wrongly claim that Dr. Hoover's plans are untethered to Defendants' conduct because they include comprehensive mental health services. That argument ignores national school mental health standards and implementation science, both of which require a comprehensive approach. *See* Ex. 1099 ¶ 3 (multi-tiered system supports ("MTSS") are "tailored" to meet the "distinct behavioral, emotional, and cognitive risks introduced by social media and the harm to school districts and the school environment"); *id.* ¶ 67 (MTSS are supported by "public health implementation science, and evidence-informed school mental health frameworks" and "peer-reviewed studies"). Moreover, Dr. Hoover directly links her plans to Defendants' misconduct. She explains: "Educators report struggling to maintain engagement and classroom cohesion amid pervasive digital distractions, while students themselves are contending with anxiety, depression, negative body image, and sleep deprivation—all caused or exacerbated by social media use. These individual-level challenges converge to strain school-wide resources, forcing districts to divert limited staffing and funding to address the mental health and behavioral consequences of digital overuse." Ex. 1000 ¶ 83; *see also* Ex. 1091 ¶ 4; Ex. 1092 ¶ 4; Ex. 1093 ¶ 4; Ex. 1095 ¶ 4; Ex. 1096 ¶ 4.

Defendants' critique of the plans' 15-year duration fares no better. The plans target harm to the Districts, not individual students, and are designed to prevent and mitigate systemic harm over time. *See, e.g.*, Ex. 1000 ¶ 6. Implementation science supports the 15-year timeline as the "necessary duration for implementing a comprehensive, district-led strategic plan . . . [allowing] for the full cycle of planning, implementation, evaluation, and sustainability." Ex. 1000 ¶ 7; *see* Ex. 1099 ¶ 74. As Dr. Hoover explains, the harms "pervade the school," and "it [is] essential to put services in place for those students who may not have yet used a platform directly themselves." Ex. 1097B at 623:10–624:4; *id.* at 627:2–628:7.

Defendants' "wish-list" characterization is equally baseless. Dr. Hoover makes clear that her staffing recommendations respond to new and exacerbated harms created by Defendants' platforms, not

to historic understaffing. Ex. 1103 ¶ 40 (plans include "additional staffing . . . necessary to respond to the mental health and learning impacts of social media"), ¶ 49 (Defendants' platforms "have not replaced historical challenges; they have created new challenges and have magnified previous challenges"). Her recommendations are a separate MTSS framework targeted specifically to social-media-related harms. *See, e.g.*, Ex. 1093 ¶ 7 ("Existing staffing and programming . . . are insufficient to address the complex impacts of social media. A comprehensive and sustainable plan requires new, dedicated staffing and professional development."), ¶ 27 ("My recommendations are additive, not duplicative . . . targeted toward addressing harms that existing efforts were never designed to address, including those stemming from compulsive platform use."), ¶ 40 (current staff face "role dilution" because of new harms from Defendants' platforms; "the solution is . . . to expand capacity").

Ultimately, Defendants' objections raise factual disputes appropriate for cross-examination, not summary judgment. *See Fuller*, 865 F.3d at 1161.

3. The Districts' harms are not derivative.

As this Court has already held, the Districts seek recovery for harms they have themselves suffered and continue to suffer due to Defendants' platforms—not for harms suffered by students. *See* MTD Order, 754 F. Supp. 3d at 968 (rejecting Defendants' derivative-injury argument and distinguishing *Ass'n of Washington Pub. Hosp. Dists. v. Philip Morris Inc.*, 241 F.3d 696, 701 (9th Cir. 2001)). Defendants' contrary claim mischaracterizes Dr. Hoover's strategic plans, which directly address the Districts' own operational and financial injuries. *See, e.g.*, Ex. 1000 ¶ 83 (explaining that student-level challenges "converge to strain school-wide resources, forcing districts to divert limiting staffing and funding to address the mental health and behavioral consequences of digital overuse.").

That Dr. Hoover's plans include student mental-health services does not render the Districts' harms derivative. Schools are legally and financially responsible for providing those services, and Defendants' platforms have magnified those burdens. *See id.* ¶ 5. As Dr. Hoover explains, "mental health challenges triggered and exacerbated by social media use have direct implications for schools, as they affect students' academic performance, social skills, and overall well-being." *Id.* ¶ 85. Defendants ignore that these are institutional harms: disruptions to school operations, depletion of budgets, and deterioration of the educational environment. *See id.* ¶¶ 83–85 (describing the "profound and far-reaching impact … on

the school environment" and the "considerable resource strain on schools").

In sum, Dr. Hoover's strategic plans are aimed squarely at remedying harms suffered by the Districts themselves, not vicariously at student injuries. As this Court has already held, the Districts "seek recovery of unique damages" they have directly sustained and continue to sustain as a result of Defendants' misconduct. *See* MTD Order, 754 F. Supp. 3d at 968.

IV. CONCLUSION

For the reasons set forth in this omnibus filing, and the accompanying briefs filed by each District, summary judgment should be denied. Genuine issues of material fact exist as to Defendants' negligence and their creation and maintenance of a public nuisance that burdens schools and communities. A jury is entitled to see and hear the evidence of Defendants' misconduct.

Respectfully submitted,

DATED: November 7, 2025

By: */s/ Previn Warren*

PREVIN WARREN
MOTLEY RICE LLC
401 9th Street NW Suite 630
Washington DC 20004
Telephone: 202-386-9610
pwarren@motleyrice.com

LEXI J. HAZAM
LIEFF CABRASER HEIMANN & BERNSTEIN, LLP
275 BATTERY STREET, 29TH FLOOR
SAN FRANCISCO, CA 94111-3339
Telephone: 415-956-1000
lhazam@lchb.com

Co-Lead Counsel

CHRISTOPHER A. SEEGER
SEEGER WEISS, LLP
55 CHALLENGER ROAD, 6TH FLOOR
RIDGEFIELD PARK, NJ 07660
Telephone: 973-639-9100
cseeger@seegerweiss.com

Counsel to Co-Lead Counsel and Settlement Counsel

JENNIE LEE ANDERSON
ANDRUS ANDERSON, LLP
155 MONTGOMERY STREET, SUITE 900
SAN FRANCISCO, CA 94104
Telephone: 415-986-1400
jennie@andrusanderson.com

Liaison Counsel and Ombudsperson

MATTHEW BERGMAN
SOCIAL MEDIA VICTIMS LAW CENTER
821 SECOND AVENUE, SUITE 2100
SEATTLE, WA 98104
Telephone: 206-741-4862
matt@socialmediavictims.org

JAMES J. BILSBORROW
WEITZ & LUXENBERG, PC
700 BROADWAY
NEW YORK, NY 10003
Telephone: 212-558-5500
jbilsborrow@weitzlux.com

ELLYN HURD
SIMMONS HANLY CONROY, LLC
112 MADISON AVE, 7TH FLOOR
NEW YORK, NY 10016
Telephone: 212-257-8482
ehurd@simmonsfirm.com

ANDRE MURA
GIBBS MURA, A LAW GROUP
1111 BROADWAY, SUITE 2100
OAKLAND, CA 94607
Telephone: 510-350-9717
amm@classlawgroup.com

MICHAEL M. WEINKOWITZ
LEVIN SEDRAN & BERMAN, LLP
510 WALNUT STREET
SUITE 500
PHILADELPHIA, PA 19106
Telephone: 215-592-1500
mweinkowitz@lfsbalw.com

MELISSA YEATES
KESSLER TOPAZ MELTZER & CHECK LLP
280 KING OF PRUSSIA ROAD
RADNOR, PA 19087
Telephone: 610-667-7706
myeates@ktmc.com

Plaintiffs' Steering Committee Leadership

RON AUSTIN
RON AUSTIN LAW
400 MANHATTAN BLVD.
HARVEY, LA 70058
Telephone: 504-227–8100
raustin@ronaustinlaw.com

AELISH M. BAIG
ROBBINS GELLER RUDMAN & DOWD LLP
1 MONTGOMERY STREET, #1800
SAN FRANCISCO, CA 94104
Telephone: 415-288-4545
AelishB@rgrd.com

PAIGE BOLDT
ANAPOL WEISS
130 N. 18TH STREET, #1600
PHILADELPHIA, PA 19103
Telephone: 215-929-8822
pboldt@anapolweiss.com

THOMAS P. CARTMELL
WAGSTAFF & CARTMELL LLP
4740 Grand Avenue, Suite 300
Kansas City, MO 64112
Telephone: 816-701-1100
tcartmell@wcllp.com

FELICIA CRAICK
KELLER ROHRBACK LLP
1201 THIRD AVENUE, SUITE 3400
SEATTLE, WA 98101
Telephone: 206-623-1900
fcraick@kellerrohrback.com

SARAH EMERY
HENDY JOHNSON VAUGHN EMERY PSC
600 WEST MAIN STREET, SUITE 100
LOUISVILLE, KY 40202

Telephone: 859-600-6725
semery@justicestartshere.com

KIRK GOZA
GOZA HONNOLD
9500 NALL AVE. #400
OVERLAND PARK, KS 66207
Telephone: 913-412-2964
Kgoza @gohonlaw.com

RONALD E. JOHNSON, JR.
HENDY JOHNSON VAUGHN EMERY PSC
600 WEST MAIN STREET, SUITE 100
LOUISVILLE, KY 40202
Telephone: 859-578-4444
rjohnson@justicestartshere.com

MATTHEW P. LEGG
BROCKSTEDT MANDALAS FEDERICO, LLC
2850 QUARRY LAKE DRIVE, SUITE 220
BALTIMORE, MD 21209
Telephone: 410-421-7777
mlegg@lawbmf.com

SIN-TING MARY LIU
AYLSTOCK WITKIN KREIS & OVERHOLTZ, PLLC
17 EAST MAIN STREET, SUITE 200
PENSACOLA, FL 32502
Telephone: 510-698-9566
mliu@awkolaw.com

JAMES MARSH
MARSH LAW FIRM PLLC
31 HUDSON YARDS, 11TH FLOOR
NEW YORK, NY 10001-2170
Telephone: 212-372-3030
jamesmarsh@marshlaw.com

JOSEPH H. MELTER
KESSLER TOPAZ MELTZER & CHECK LLP
280 KING OF PRUSSIA ROAD
RADNOR, PA 19087
Telephone: 610-667-7706
jmeltzer@ktmc.com

HILLARY NAPPI
AWK ATTORNEYS

1133 WESTCHESTER AVE, SUITE N-224
WHITE PLAINS, NY 10604
Telephone: 914-468-4840
hnappi@awk-saa.com

EMMIE PAULOS
LEVIN PAPANTONIO RAFFERTY
316 SOUTH BAYLEN STREET, SUITE 600
PENSACOLA, FL 32502
Telephone: 850-435-7107
epaulos@levinlaw.com

RUTH THI RIZKALLA
THE CARLSON LAW FIRM, PC
1500 ROSECRANS AVE., STE. 500
MANHATTAN BEACH, CA 90266
Telephone: 415-308-1915
rrizkalla@carlsonattorneys.com

ROLAND TELLIS
DAVID FERNANDES
BARON & BUDD, P.C.
15910 Ventura Boulevard, Suite 1600
Encino, CA 91436
Telephone: 818-839-2333
rtellis@baronbudd.com
dfernandes@baronbudd.com

DIANDRA "FU" DEBROSSE ZIMMERMANN
DICELLO LEVITT
505 20th St North
Suite 1500
Birmingham, Alabama 35203
Telephone: 205-855-5700
fu@dicellolevitt.com

Plaintiffs' Steering Committee Membership

JOSEPH VANZANDT
BEASLEY ALLEN
234 COMMERCE STREET
MONTGOMERY, AL 36103
Telephone: 334-269-2343
joseph.vanzandt@beasleyallen.com

Federal/State Liaison

Attorneys for Individual Plaintiffs

ATTESTATION PURSUANT TO CASE MANAGEMENT ORDER NO. 27

I, Previn Warren, attest that the evidence cited herein fairly and accurately supports the facts as asserted.

Dated: November 7, 2025 By: */s/ Previn Warren*

FILER'S ATTESTATION

I, Previn Warren, hereby attest, pursuant to N.D. Cal. Civil L.R. 5-1, that the concurrence to the filing of this document has been obtained from each signatory hereto.

Dated: November 7, 2025

By: */s/ Previn Warren*